

Disorder Identity: Opportunity or Obstacle?

by

Hugh Alistair Campbell Webb

June 2017

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

Clinical Psychology

of The Australian National University

© Copyright by Hugh Alistair Campbell Webb 2017

All Rights Reserved

Statement of Originality

I, Hugh Alistair Campbell Webb, declare that the text and work presented in this thesis is my own and that no sources other than those mentioned in the text and its references have been used in creating it.

A handwritten signature in black ink, appearing to read 'H. A. Campbell Webb', written in a cursive style.

Dedication

To Dunc.

Acknowledgments

I would like to thank my supervisor, Professor Michael Platow. Michael was the first person to challenge my physics envy and persuade me to take social processes seriously. Throughout my studies he has encouraged me to chase big ideas and provided some essential intellectual tools to help pin these ideas down, chop them up and lump them together again. Without Michael's support and patience I would not have been able to begin, let alone complete, this thesis.

I also want to thank Dunc for walking alongside me as I tried to turn a series of thought fragments into a sustained argument. Nobody understands the concerns I want to point to in this thesis better than him.

And finally, thank you for the love and patience shown by the rest of my family: my wife Megan, Mum, Sal and Ian, Anne, Adrian, Jenny, Rachel and Kev. You have all contributed in big and little ways to make this thesis possible. Thank you.

Abstract

This thesis develops a social psychological analysis of Hacking's (1999) "looping kinds" account of psychopathology. Informed by Social Identity Theory (SIT, Tajfel & Turner, 1979) and Self Categorization Theory (SCT, Turner et al., 1987), a Disorder Identity Analysis is proposed, according to which the symptoms of psychopathology are partly shaped by the interaction between expert clinical categorisation schemes and first person Disorder Identities, represented as self-categories and shared by fellow individuals who experience disorder. A set of theoretically informed predictions are then derived for empirically scrutinising Hacking's "looping kinds" account and to help explain apparent coherence, stability and change observed in psychopathology. Results from five experimental studies are then reported, demonstrating support for one of the predictions derived from the Disorder Identity analysis: that identifying with fellow disorder sufferers could either enable or be a barrier to clinical change, depending on the group's normative understanding about the malleability of the condition. These findings suggest that the proposed Disorder Identity Analysis shows promise as a means of specifying, and empirically operationalising, Hacking's "looping kind" arguments.

Table of Contents

List of Tables	vi
List of Figures.....	vii
Preface	1
Chapter 1: What is Mental Disorder?.....	4
Where is mental disorder?	4
Explaining coherence, stability and change: the Disease Model.....	5
The conceptual problem of psychopathology	5
The Disease Model.....	8
Implications of the disease model.	15
Criticisms of the Disease Model	18
Socio-political critiques.	18
Engel’s bio-psycho-social critique.	20
Empirical critique.	21
Homeostatic Property Clusters—an alternative to essentialist natural kinds?	23
Boyd’s Homeostatic Property Cluster account.	23
Kendler et al.’s HPC account of mental illness.....	30
Criticisms of HPC internalism	32
Arbitrariness objection to HPC internalism.	33
Empirical inadequacy objection to HPC internalism.	34
Testing the “looping” hypothesis: social identity as an HPC causal mechanism	38
Summary: towards a positive alternative HPC account	50
Chapter 2: The Disorder Identity Analysis: the Disorder Identity and HPC	
Frameworks and Recovery from Mental Disorder	58
Recovery as cure, outcome or orientation?.....	60

Recovery-as-an-orientation.....	62
Recovery, adaptation and the self	66
Defining Disorder Identity	67
The consequences of Disorder Identities on HPC categories	74
Disorder Identities in action.....	79
Recovery: summary and conclusions	86
Chapter 3: Social-Psychological Processes in Mental Disorder	91
Internalising illness labels: labelling theories, self-stigma and stereotype threat	92
Labelling Theory	92
Modified Labelling Theory	95
Self-Stigma and Stereotype Threat.....	103
Summary and implications of Labelling Theory research.....	106
Content of disorder categories: essentialism, malleability and motivation to	
change	107
Lay psychiatric theorising.....	108
Self-theories and achievement motivation.....	114
Attribution and motivation as a dynamic, group process	118
Summary and implications	122
Chapter 4: A Social Identity Analysis of Disorder and Recovery	125
The Social Identity Approach	126
Social Identity Theory and intergroup behaviour	126
Self-Categorization Theory.....	131
Social identity analysis of social influence.....	136
Applying the Social Identity Approach to understand the dynamics of Disorder	
Identities.....	139

Disorder Identities, context dependent self-stereotyping and symptom change.	142
Disorder Identities as a means to adaptively modify the social environment.....	145
Using Disorder Identity Adaption Strategies to modify the self.....	151
Disorder Identity Adaption Strategies as heterostatic mechanisms	155
Disorder Identity Adaption Strategies as heterodynamic mechanisms	160
Summary: Disorder Identity Adaption Strategies as a dynamic system	167
Chapter 5: Introduction to the Empirical Chapters	175
Chapter 6: Identity Processes as Determinants of Individual Malleability Beliefs	189
Study 1 introduction	197
Study 1 method	197
Participants and design	197
Materials and procedure.....	198
Study 1 results.....	202
Preliminary analyses and manipulation checks.....	202
Main analysis	203
Study 1 discussion.....	206
Study 2 introduction	207
Study 2 method	208
Participants and design	208
Materials and procedure.....	208
Study 2 results.....	211
Manipulation checks and preliminary analyses	211
Study 2 main analysis	212
Follow-Up conditional mediation analysis	214
Study 2 discussion.....	219

Study 3 introduction	222
Study 3 method	222
Participants and design	222
Materials and procedure.....	223
Study 3 Results.....	224
Manipulation checks and preliminary analyses	224
Study 3 main analysis	226
Study 3 discussion.....	228
General discussion.....	229
 Chapter 7: Study 4 – The Interaction Between a Malleable Group Norm and	
Disorder Identification Predicts Reductions in Symptomatology Over	
Time.	234
Study 4 method	238
Participants and procedure.....	238
Measures	238
Study 4 results.....	240
Preliminary analyses.....	240
Main analysis	243
Study 4 discussion.....	246
 Chapter 8: Study 5: The Interaction Between Disorder Identity and the Group	
Malleability Norm Predicts Personal Malleability Beliefs and Lower	
Anxiety Symptoms.	250
Study 5 method	253
Participants and design	253
Materials and procedure.....	254

Study 5 results.....	256
Measures and manipulation	256
Main analysis	258
Moderated mediation analysis.....	261
Study 5 discussion.....	267
Chapter 9: Conclusion	270
Framing the issue: What is mental disorder and where is it?	271
Building the conceptual foundations of the disorder identity account.....	276
Social psychological mechanisms for explaining Disorder Identity dynamics ..	279
A critical test of the Disorder Identity Analysis: the collective mobility thesis .	283
Wider implications of the Disorder Identity Analysis	288

List of Tables

Table 1: <i>The HPC framework: homeostatic, homeodynamic, heterostatic, and heterodynamic mechanisms</i>	76
Table 2: <i>Disorder Identity Adaption Strategies to modify the social environment</i>	150
Table 3: <i>Four kinds of heterostatic (stabilising) norms</i>	159
Table 4: <i>Four kinds of heterodynamic (transitional) norms</i>	161
Table 5: <i>Disorder Identity Framework: a taxonomy of Disorder Identity Adaption Strategies</i>	173
Table 6: <i>HPC Framework: Disorder Identities as HPC mechanisms</i>	174
Table 7: <i>Items Measuring Social Identification and Individual Fixed Beliefs</i>	201
Table 8: <i>Study 1 means, standard deviations and bivariate correlations between main study variables</i>	204
Table 9: <i>Study 2, means, standard deviations, and bivariate correlations between main study variables</i>	212
Table 10: <i>Study 3, means, standard deviations, and bivariate correlations between main study variables</i>	226
Table 11: <i>Study 4, means, standard deviations and bivariate correlations between main study variables</i>	242
Table 12: <i>Study 5 means, standard deviations and bivariate correlations between main study variables</i>	259
Table 13: <i>Summary of studies testing the Collective Mobility Thesis</i>	291

List of Figures

Figure 1: Conceptual distinction between internalism and externalism.	10
Figure 2: The disease model and two possible HPC models.....	25
Figure 3: Parallels between HPCs (Boyd, 1991) and constraint relation models of categorization (McGarty, 1999)	28
Figure 4: A simplified hypothetical model of the social and cultural determinants of the genetically determined disease Phenylketonuria.....	35
Figure 5: Causal models of mental illness: a two dimensional conceptual framework.	52
Figure 6: The Disorder Identity Analysis of looping effects: the interaction between the Disorder Identity Framework and the HPC Framework.	70
Figure 7: Interaction between perceiver readiness and fit determines category salience and its consequences.....	135
Figure 8a: “Fixed” word cloud	200
Figure 8b: “Malleable” word cloud	200
Figure 9: Significant two-way interaction between shy group identification and manipulated malleability norm on personal fixed beliefs at Time 2 (controlling for Time 1 fixed beliefs).	205
Figure 10: Conditional mediation model comparing social influence and social differentiation explanations.	218
Figure 11: Shy identification and a fixed group norm (at Time 1), interact to predict reductions in Social Interaction Anxiety at Time 2.....	245
Figure 12: Conditional mediation model assessing moderation of the fixed group norm by group identification on Time 2 Individual Fixed Beliefs and subsequent relationship to Self-rated State Anxiety (controlling for the effect of Time 1 Fixed Beliefs and the interaction between Time 1 Fixed Beliefs and Identification).....	265

Figure 13: Conditional mediation model assessing moderation of the fixed group norm by group identification on Time 2 Individual Fixed Beliefs and subsequent relationship to Partner Ratings of Anxiety (controlling for the effect of Time 1 Fixed Beliefs and the interaction between Time 1 Fixed Beliefs and Identification).....	266
--	-----

Preface

In this thesis I examine whether social identification with mental disorder categories could help explain coherence, stability and change observed in mental disorder. It might be assumed, given the extraordinary breakthroughs in neuroscience and genetics, that there is simply no need for such coarse-grained social explanations of mental disorder. Yet on the contrary, many within the discipline of psychiatry and clinical psychology, and from radically diverse theoretical perspectives, have in recent years come to the conclusion that there is nothing less than a crisis of confidence in psychiatric categories (Frances, 2013; Insel et al., 2010; Pemberton & Wainwright, 2014a). While many suggest this crisis will not be solved by merely looking more carefully inside individual, mentally disordered heads, others fear that invoking social explanations will only add to the conceptual confusion. I argue that a social identity based account of disorder, informed by Social Identity Theory (SIT, Tajfel & Turner, 1979) and Self Categorization Theory (SCT, Turner et al., 1987), could help put social processes on a more rigorous theoretical and empirical footing, and so generate fresh and much needed insights into the nature of mental disorder and recovery.

To undertake such a task requires a substantial shift in thinking and careful attention to the meta-theoretical assumptions that underpin empirical approaches to understanding mental disorder. In the first chapter of the thesis, I therefore start by explaining the central conceptual problem of accounting for apparent stability and change observed in mental disorder. I survey different theoretical approaches to this problem and show how these can be mapped on to two main dimensions. The first dimension relates to the question of *what* is mental disorder: should it be understood as an essence or as a set of mechanisms? The second dimension of the problem relates to the question of *where* mental disorder lies: should it be understood as existing solely inside individual heads, or should disorder be understood as, in some sense, being caused by entities or processes outside of, or in between, individual

heads? I argue that understanding mental disorder as partly defined by social group mechanisms offers important and novel insights into the nature of disorder. Most importantly, I argue that a Disorder Identity Analysis could transform this from a provocative suggestion into a series of testable empirical predictions.

In the second chapter, I build the conceptual foundations for the Disorder Identity Analysis by elaborating what I call Disorder Identities and Disorder Identity Adaption Strategies. I further argue that the notion of Disorder Identity has important connections with the concept of recovery-as-an-orientation. By showing that Disorder Identities are what enable groups of people with disorder to collectively modify themselves and their social environment, I demonstrate the extraordinarily diverse ways in which people with disorder can “recover”. Importantly, I argue that these Disorder Identities are informed by lay theories which often do not align with expert and scientific accounts of disorder, but which nevertheless have causal consequences that can systematically change the properties of disorder categories. Specifically, I outline the central mechanisms that are needed to explain how Disorder Identities could change the stock and flow of people who meet criteria for disorder, as well as the average properties of those with disorder.

In the third and fourth chapters I elaborate the social-psychological processes that could underpin these postulated Disorder Identity mechanisms. In the third chapter I address some of the challenges posed by certain social-psychological analyses of disorder, and review the literature showing the role that lay understandings of disorder play in shaping symptom perception and motivation. In the fourth chapter, I finally elaborate in some detail the social identity based psychological mechanisms that could explain coherence, stability and change in mental disorder categories.

However, in developing the Disorder Identity Analysis I uncover a theoretical ambiguity within SIT about whether transitioning out of a low status group, such as a

disorder category, is necessarily individuating, or can in some sense be understood as a collective process. This very specific question, derived from the Disorder Identity framework, forms the basis of the subsequent empirical chapters.

Across five empirical studies, described in Chapters 6 to 8, I test the proposition that identifying with fellow disorder sufferers could either enable or be a barrier to clinical change, depending on the group's normative understanding about the malleability of the condition. I show how these findings not only begin to clarify the social-psychological processes that enable collective transition out of a group, but also demonstrate the feasibility of the Disorder Identity Analysis as an empirical project. In Chapter 9 I summarise the core insights developed in the Disorder Identity Analysis, and reflect on the wider applications, implications and possible extensions of the analysis.

One final point to note is that throughout the thesis I try to use the term “mental disorder”, even though, as I shall make clear in Chapter 1, in many ways this prejudices the conceptual nature of the subject matter.

Chapter 1: What is Mental Disorder?

Where is mental disorder?

A very common, some would say trivial, proposition is that mental disorders are disease entities that exist inside individual brains. However, in recent years, the evidence to support this simple proposition cannot be found for any mental disorder. As anybody who has experienced or supported a person with mental disorder can attest, this cannot mean that mental disorder is not real. Perhaps researchers have simply not looked hard enough—after all, it is notoriously difficult to prove an entity’s non-existence. In this chapter, I argue that part of the reason mental disorders cannot be found inside individual heads may be that we are looking in the wrong place and that we have misconstrued the role of social processes in the conceptual underpinnings of disorder. Indeed despite furious agreement, from diverse theoretical perspectives, that social processes must, in some way, influence mental disorder, I argue that beneath this apparent agreement rests an invariably impoverished understanding of the social.

To redress this, I argue that Hacking’s (1999) notion of “looping kinds”—whereby groups of people with disorder collectively react to their disorder label and so change the nature of disorder categories—provides a more compelling and dynamic model for understanding mental disorder. On this view, mental disorder needs to be understood less like a specimen in a cabinet and more like a species evolving in an ecosystem, one that actively shapes and is shaped by both its genetic and social-environmental context. Perhaps most importantly, I further claim that SIT and SCT provide a theoretical framework for making these dynamic processes empirically tractable.

This chapter is divided into five parts. In the first part I outline the central conceptual challenge of explaining coherence, stability and change in mental disorder. I then examine the preeminent theoretical account for explaining this coherence, stability and change: the so-

called disease model of mental illness and five, inter-related tenets that underlie it. This is important, because many of the subsequent debates hinge on one or more of these tenets. In the second part I outline various critiques of the disease model with a particular focus on the failure to find disease essences to explain psychiatric disorder. In the third part, I present an increasingly popular alternative that avoids several problems with essentialism, the so-called Homeostatic Property Cluster (HPC) account developed by Boyd (1991), and discuss an influential application of this model to mental disorder that has recently been developed by Kendler, Zachar, and Craver (2011). In the fourth section I show that while understanding mental disorder in terms of a series of mechanisms, rather than essences, is an important advance, it is problematic in part because it does not contemplate the possibility that some of these mechanisms could be group-based processes that operate not just within, but between people with disorder. In the fifth and final section of the chapter, I provide a conceptual overview of the various meta-theoretical approaches to disorder that have been reviewed and summarise the conceptual assumptions that will underpin the analysis of Disorder Identity that I develop in later chapters.

Explaining coherence, stability and change: the Disease Model

The conceptual problem of psychopathology

A core conceptual problem of psychopathology is explaining both stability and variation in mental illness. On one hand, very similar syndromes of putatively psychiatric behaviour seem to occur across time and space. Individuals exhibiting patterns of behaviour resembling depression, mania, psychosis, and obsessive compulsive disorder are recorded in vastly different historical contexts: in the writings of ancient Babylonians (Kinnier Wilson, 1967), Greeks (Jeste, del Carmen, Lohr, & Wyatt, 1985), Egyptians (Nasser, 1987), and several ancient eastern cultures (Harris, 2013). Contemporary evidence from very large, cross-national epidemiological studies also finds substantial cross-cultural commonalities in

schizophrenia (Saha, Chant, Welham, & McGrath, 2005), major depression (Ferrari et al., 2013), anxiety disorders (Baxter, Scott, Vos, & Whiteford, 2013), and bipolar disorder (Weissman et al., 1996).

But on the other hand, there is also substantial cross-cultural variation in the kind, prevalence, patterning and correlates of psycho-pathology. There are very well documented psychiatric syndromes that seem to occur only in highly localised socio-cultural and historical niches. The glass delusion, for instance, emerged among the European elite in the middle ages and caused afflicted individuals to become obsessed with the idea they were made of glass and so went to extraordinary lengths to avoid physically shattering. The condition was prevalent enough to be described by Descartes (2013 / 1641) and was included in Burton's *Anatomy of Melancholy* (2012 / 1621). But since the seventeenth century, no confirmed cases have been reported (Speak, 1990).

More contemporary evidence shows that several psychiatric syndromes occur only in certain traditional societies, such as *koro*, *pibloktoq*, *susto* and *amok* (Yap, 1974), whereas other conditions are far more prevalent in developed countries, including dissociative identity disorder, pre-menstrual dysphoric disorder, and anorexia nervosa (Chrisler & Caplan, 2002). As these cultures change, so too do patterns of putative psycho-pathology. For instance, the prevalence of anorexia nervosa within developing countries increases with increasing exposure to western media (Becker, Burwell, Herzog, Hamburg, & Gilman, 2002; Eddy, Hennessey, & Thompson-Brenner, 2007). Some kinds of psychopathology even appear to be socially contagious (Joiner & Katz, 1999), sometimes rapidly so. Mass motor hysteria is very widely documented and can cause affected individuals to experience convulsions, tremors and paralysis, can spread line-of-sight within tight-knit communities, and seems to emerge and resolve within days to weeks (Bartholomew & Wessely, 2002).

Many clinical disorders display elements of longitudinal and cross-cultural stability, but also distinct variation. For example, while Post-Traumatic Stress Disorder (PTSD)-like syndromes have been observed across disparate cultures (Osterman & De Jong, 2007), their presentation varies. American and Cambodian sufferers have similar fear reactions in PTSD, but the interpretation of these reactions and their impact on functioning is vastly different (McNally, 2012). Even within cultures, longitudinal variation can be observed. It might be assumed, for example, that the syndrome of “shell-shock”, experienced by thousands of soldiers during World War I and World War II, must be symptomatically identical to present day PTSD. But detailed analysis of British combat veterans’ medical records since the Boer War finds that while experiencing flashbacks was a common and very prominent symptom after the first Gulf War, the same symptom was observed in only around one percent of World War I and II veterans, and was never reported among Boer War veterans (McNally, 2012).

The issues are further complicated by the difficulties of attempting to corral and stabilise the measurement and definition of putative disorders. Substantively different disorders can be wrongly assimilated, while differing idioms of distress can obscure deeper commonalities. Even within a culture, the definitional issues are legendary. When the first Diagnostic and Statistical Manual was released in 1952 there were 26 disorders, but this had increased to 374 by the fourth (1994) and fifth editions (2012). New disorders were described, some disbanded, some merged, some split and some redefined.

The obvious question to ask is the extent to which this variation reflects “real” changes in human psycho-pathology or whether these changes are an artefact. Logically, if these differences are real then this could be because of genetic drift or *de novo* genetic mutations (Neale et al., 2012); because of novel psychopathogenic infectious agents (Foster & Neufeld, 2013); or because varying social and environmental circumstances have triggered

varying psychological and behavioural responses. To the extent observed variation is an “artefact”, then this could be: because disorders have been mis-described or masked by superficial social and linguistic differences; because of varying rates of diagnosis and treatment; because some forms of formerly pathological behaviour are no longer considered socially problematic and so are no longer *called* a disorder; or, conversely, because thresholds for defining disorder have been lowered, or broadened, to encompass behaviours that were not previously considered pathological (Haslam, 2016)

To disentangle these competing possibilities therefore requires a robust theoretical account of what mental disorders are. But in recent years, prominent researchers, from diverse perspectives within psychiatry and clinical psychology, seem to agree on little other than that there is a crisis of confidence in diagnostic categories (Casey et al., 2013; Frances, 2013; Insel et al., 2010). This crisis has spurred a renewed interest in examining different meta-theoretical positions for understanding the nature of mental disorder. These positions, in turn, direct what mental illness is, where to look for it, and how to know it when one sees it.

The Disease Model

An important starting point for many discussions about the nature of mental illness is the disease model (also sometimes called the bio-medical model). This model reflects a number of interrelated propositions that have been canvassed in several reviews (Beebe & Sabbarton-Leary, 2010; Haslam, 2002b; Kendler, 2005b; Zachar & Kendler, 2007). Helpfully, the disease model foregrounds several conceptual fissures that divide competing explanatory frameworks.

Essentialism. Perhaps the most important tenet of the disease model is that mental illness is an essentialist natural kind. In the context of the classic disease model, this is the claim that all mental illnesses have some defining essence (e.g., deletions on chromosome 7

in William's disease) that all and only sufferers of that condition possess. According to this essentialist natural kind view, the externally observable signs and symptoms of an illness are caused by these underlying essences and so illnesses exist independently of human systems of classification. The natural kind view contrasts with the so called nominalist view of mental illness, according to which specific mental "illnesses" are mere labels for syndromes, that is, convenient or conventional ways of organising signs and symptoms that do not necessarily imply anything about the existence of an underlying basis to the condition.

The putative advantage of the natural kind view over the nominalist view is that by accurately mirroring the entity that generates surface-level similarities, clinicians and researchers cannot just understand why signs and symptoms co-occur, they can make accurate predictions about them. It is this feature of natural kinds that is so generative in some other scientific disciplines. Knowing that gold is essentially made up of atoms with 79 protons does not just allow one to reliably identify a novel instance of gold (and be sure it is not similar-looking iron pyrites), but it allows one to make law like predictions about how that novel instance will behave under particular circumstances (melt at 1064 degrees centigrade, conduct electricity). This "inductive potency" explains why essentialist accounts are, in theory, so powerful. If mental disorders are natural kinds then an accurate diagnosis is richly informative—enabling strong inferences about aetiology, pathogenesis, and treatment. If mental disorders are natural kinds then it would be quite possible for somebody to be mentally ill but asymptomatic, and therefore potentially amenable to interventions to slow or prevent symptomatic sequelae. By contrast, if mental disorders are merely descriptive syndromes, then asymptomatic mental illness is logically impossible.

Internalism. The second, related, tenet of the disease model is that mental illness is primarily, if not exclusively, caused by processes within the brain such as faulty neurotransmitters or a neurotic personality. According to this so-called *internalist* view, while

mental illness may have various external social and interpersonal correlates, these are largely epiphenomenal in the aetiology of mental illness. The internalist view contrasts with the so called externalist theory of mind according to which events outside the brain and body play a dominant causal role in explaining mental states. For a strong externalist, mental illness is to a brain as a bad TV program is to a TV set. The strong externalist concedes that brain states are proximal causes of mental states, in a trivial sense of transmitting the effects of external influences into thoughts and behaviours. But the strong externalist downplays the importance of internal states causing mental illness, over and above the mediated effects of phenomena external to the brain and body (the physical environment, interpersonal relationships, social institutions). The internalism / externalism distinction is depicted in Figure 1. Of course, the externalist and internalist positions are not strictly exclusive. Few would deny that internal and external causes mutually influence mental illness. Many *would* dispute the *extent* to which internal or external causes should be privileged in our understanding of mental illness (Chanock et al., 2007; Cuthbert & Insel, 2013; Kandel, 1998; Kendler, 2005b; Wilson, 2004).

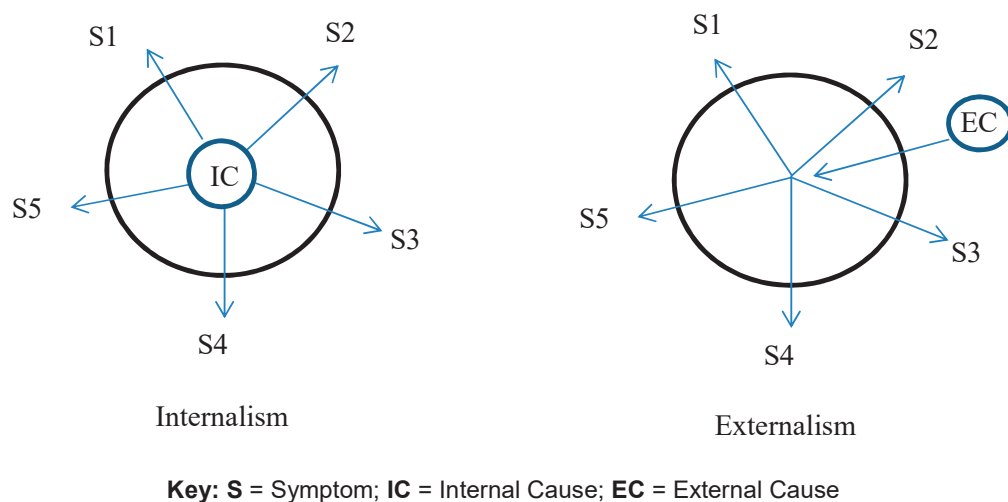


Figure 1: Conceptual distinction between internalism and externalism.

Reductionism. The third tenet of the disease model is that mental phenomena can best be understood *reductively*. In other words, we should privilege explanations and methodologies that focus on microphysical entities (genes, neuro-transmitters) rather than the macroscopic entities (individuals, families, societies) they constitute. Again, faith in methodological reductionism is a matter of degree. For some, so-called “ruthless”, reductionists (e.g. Bickle, 2006; see also Guze, 1989) it is not possible for explanations of mental illness to be too biological (except perhaps if they can be further reduced to fundamental physics). They argue that explaining mental phenomena in terms of relatively coarse-grained entities, like cognitive processes or social groups, risks over-determining what is better explained by fundamental molecular processes on which all higher-levels must depend (Cromwell & Panksepp, 2011; Kandel, 1998). Reductionists therefore argue that denying methodological reductionism entails either redundancy or rejecting materialism (roughly, the idea that the only things that exist are the physical entities described by physics).

In the context of mental illness research, a reductionist methodology is partly a consequence of the first two tenets described previously. If what the disease model tries to explain is the disease-causing entity within the brain, then it follows intra-individual targets should be privileged in empirical investigations. But reductionism is not synonymous with internalism or essentialism. Reductionists would have little truck, for example, with the notion that a neurotic personality is causally relevant to mental illness (even though it is internal) unless neurotic personality can be further explained neuro-biologically. Conversely, it is conceivable that an externalist could endorse reductionism—for example, by exploring how molecular phenomena in the environment influence mental states (also described at a molecular level). Enthusiasts of such an approach have suggested attempting to map the “exposome”: a detailed, longitudinal model of the environment that aspires to the same level

of precision as maps of the human genome (Wild, 2005). But the environment is a big place. As some have wryly noted, a molecular model of it would be a “long-term project” (Sadler, Wiggins, & Schwartz, 1994, p. 287).

These possibilities show how denying methodological reductionism is not the same as denying the ontological thesis of materialism, as is sometimes assumed (Cacioppo & Berntson, 1992; Putnam, 1980). Rather, it reflects, in part, humility about the computational constraints on reductionism. To propose that some antecedent or consequence of mental illness is best understood at a more macroscopic level does not require additions to our ontology. Nor does a denial of reductionism assume “spooky action at a distance” in which molar phenomena are unmediated by entities that can be described at a microphysical level (Craver & Bechtel, 2007). Gerard’s (1955, p. 47) memorable dictum “there is no twisted thought without a twisted molecule” is therefore trivial if taken to mean that molecules are the spatially proximal causes of mental states. Barring non-physical causation, what else could underlie mental states? *Substantive* debate only arises about whether decomposing entities into smaller components will adequately explain how, when and why brain molecules “twist”. Reductionists insist it will. Others suspect this will be as satisfactory as explaining a building’s architecture by describing the molecular properties of bricks and steel. And so, empirically plausible alternatives have been developed to ground non-reductive but *materialist* accounts of mental states (Cacioppo, Berntson, Sheridan, & McClintock, 2000; Craver & Bechtel, 2007).

Discontinuism. The fourth tenet of the disease model is that mental disorders are discrete categories, known in short as discontinuism. This is a corollary of essentialism. If illness is caused by a single, discrete entity, that all and only sufferers possess, then it follows that the presence or absence of illness is an all or nothing affair, and the boundary between illness and health is sharp. Interestingly, as Haslam (2002b, 2014) has eloquently argued, the

reverse is not true. The existence of a sharply disjunct group displaying particular signs and symptoms in a population does not guarantee the presence of a natural kind. Haslam notes that such syndromes can be generated through non-linear interactions between multiple disease-causing entities (none of which are individually necessary and sufficient), among other mechanisms. Moreover, manifestly socially determined beliefs can be discontinuous. Being a Trotskyite is one famous example (Meehl, 1995).

Nevertheless, distributions of signs and symptoms provide an important constraint on the kinds of causal mechanisms that might generate a particular condition. To accurately measure these distributions, Meehl and colleagues (1995, 1994, 1996) have developed sophisticated taxonometric procedures for determining whether a disorder displays a discontinuous (i.e. taxonic) latent structure. If an illness is not discontinuous, it may be “fuzzy” (i.e. with a non-sharp discontinuity), or smoothly continuous (although this is often artificially treated as a category on pragmatic grounds, for example by defining depression using categorical, DSM criteria—see Zachar, 2000). Taxonometric procedures therefore do not definitively prove essentialism but provide an important empirical basis for assessing its plausibility.

Ahistoricism. The fifth tenet of the disease model is that mental illnesses are largely value neutral, ahistorical and cross-culturally invariant. In part, this tenet is a consequence of essentialism. If mental illnesses are defined essentially, then their existence can be discovered empirically, rather than constructed by arbitrary human stipulation (as they are for the nominalist). Informed by essentialism, proponents of the disease model would concede there may be longitudinal and cross-cultural variation in the outward behavioural signs of mental illness, but would argue this variation is because varying environments differentially trigger disease-causing essences and because cultural variation leads to superficial differences in symptomatology caused by measurement error and reporting bias (Cheng, 2001). In

principle, they would argue, this noise could be eliminated (for example by examining biomarkers that would indicate the presence or absence of psychological disease, see Hucklenbroich, 2014). In contrast to this so-called etic view, according to the emic view, cross-cultural variation in mental illness can reflect substantive differences in underlying patho-etiology across cultures (Stier, 2013).

A further issue arises when considering whether there is some non-value-laden way of deciding whether a psychological phenotype is pathological or not (Stier, 2013). Boorse (1977) argues that pathology can be objectively defined as the failure of an internal mechanism to perform a statistically typical biological function that enables survival or reproduction (see also Hucklenbroich, 2014). There are several problems with this formulation, among them, that by these criteria homosexuality is psycho-pathology and, second, even though a biological part may “fail”, people can adapt themselves or their environment to function effectively (Cooper, 2012). More formally, others argue that Boorse falsely attempts to derive what is a question of evaluation from positive description (of our evolutionary history) and so commits the naturalistic fallacy (Moore, 1960).

Wakefield (1999) proposes an influential alternative to Boorse’s view by (more or less) retaining Boorse’s dysfunction criteria, but adding that this dysfunction must also be harmful. Wakefield concedes that defining harm is value laden, but claims that his hybrid “harmful-dysfunction” account strikes the right balance between objective and normative criteria and thereby rules out the problem of over-diagnosing. Critics point out, though, that even this hybrid formulation is still vulnerable to naturalistic fallacy objections (Houts, 2001). Undeterred, proponents of the disease model often concede that values may play a role in psychiatric classification but only in the limited sense that they are “involved in determining which psychiatric kinds deserve clinical attention” (Kendler et al., 2011, p. 120). On this view, therefore, shifting normative standards of harm and deviance will influence

what counts as mental illness, but not the bedrock essences that generate constellations of psychiatric signs and symptoms.

Implications of the disease model.

In sum, the disease model captures a series of intimately related meta-theoretical assumptions that characterise the nature of mental illness. The position can be summarised by the following “isms”: essentialism, internalism, reductionism, discontinuism, and ahistoricism. Before discussing the implications and limitations of the model, it is important to address an obvious criticism of these tenets on empirical grounds. Essentialism can be easily criticised because there are currently few if any psychiatric conditions for which a discrete, causal, biogenetic basis has been identified (Zachar & Kendler, 2007). Indeed the DSM uses explicitly nominalist criteria, warns against reifying diagnostic categories and deliberately describes these as “disorders”, not diseases, to avoid essentialist connotations.

This criticism conflates the disease model as a normative goal (i.e., how mental illness should be understood) with a positive description of psychiatric nosology, based on our current understanding. Indeed many biological psychiatrists would readily accept the criticism but deny this vitiates the disease model. For example, Grebb and Carlsson (2009) argue that because “there is little reason to believe that these (DSM) diagnostic categories are valid, in the sense that they represent discrete, biologically distinct entities” (p. 1), then it is little wonder that progress in psychiatric research has been limited relative to many other medical sciences (Kessler et al., 2005). Grebb and Carlsson echo an emerging consensus among biological psychiatrists and others (Akil et al., 2010; Cuthbert & Insel, 2013; Insel & Wang, 2010; Manolio et al., 2009) by claiming that researchers have failed to embrace the disease model vigorously enough and that brain sciences are sufficiently mature to establish a biologically-based diagnostic system (Cuthbert & Insel, 2013). Other clinicians and researchers are more pragmatic. Kendell and Jablensky (2003), for example, concede that

DSM and ICD disorders largely fail to conform to the disease model ideal, and so are formally invalid, but that such nominalist nosologies are still useful constructs for practising clinicians until such time that geneticists and neuroscientists can establish a robust, biologically-based diagnostic system (Casey et al., 2013; Jablensky & Waters, 2014).

Understanding the disease model as a normative ideal not only dispels an obvious criticism, it helps frame a very influential account of how we should correctly understand mental illness, with significant implications for the way empirical research and clinical practise should proceed. The National Institute of Mental Health research has recently established the Research Domain Criteria (RDoC), that seeks to promote a “paradigm shift” (p. 34) in psychiatric classification, informed by genetics and neurobiology. The three guiding assumptions for the criteria, explicitly and implicitly, reflect many of the disease model tenets:

First, mental illnesses are presumed to be disorders of brain circuits. Secondly, it is assumed that the tools of clinical neuroscience, including functional neuroimaging, electrophysiology, and new methods for measuring neural connections can be used to identify dysfunction in neural circuits. Third, the RDoC approach presumes that data from genetics research and clinical neuroscience will yield biosignatures that will augment clinical signs and symptoms for the purposes of clinical intervention and management. (p. 33)

The disease model ideal also implies that a necessary precondition for *genuine* recovery is a cure that directly targets a specific biological entity or process within sufferers’ brains that causes the condition. The paradigm example of this hope is the case of general paresis: a neuropsychiatric condition demonstrated in 1914 to be essentially a consequence of bacterial infection by *Treponema pallidum*, a spirochete bacterium which is now fully treatable with antibiotics (Hutto, 2001). Assuming other psychiatric conditions similarly

conform to the disease model, understanding mental illness and recovery is an empirical search for the relevant spirochetes and penicillins for each disorder.

The third implication of the disease model is that it implies a highly circumscribed role for psychological and social factors. A simplistic criticism of the disease model is that it implies biology is destiny and so social and psychological factors play no role. Proponents of the disease model can easily counter that this misunderstands the metaphysics of essentialist natural kinds. Essences are just law-like dispositions to behave in certain ways in particular contexts. Just as gold is molten, but only if its temperature exceeds 1064 degrees, a person with anorexia nervosa may be asymptomatic in some cultural contexts but would certainly display symptoms in the presence of specific environmental triggers. This is the logic of the diathesis-stress model: a latent disease-causing entity (the diathesis) is assumed to mediate the effects of environmental stressors (Cramer, Waldorp, van der Maas, & Borsboom, 2010).

But the requirement that a localised disease entity mediates illness signs is still a stringent constraint on the role of social and environmental factors. It allows for the conceptual partitioning and privileging of the disease causing essence. For example, the RDoC proposes a matrix that crosses intra-individual entities (genes, molecules, cells, circuits, physiology, observable behaviour) with specific functional systems within the brain. It then lists “developmental processes and the environment” (p. 33) as an orthogonal dimension that, for example, may influence the phenotype via epigenetic mechanisms.

The fourth implication of the disease model is that first person experience and the subjective meaning of illness for the sufferer has no direct causal import. As with social factors, these beliefs may influence treatment seeking and compliance, but they will not directly change the underlying disease essence. The perspective is therefore largely incompatible with a wide range of so-called narrative approaches (Zachar & Kendler, 2007), from both the psychoanalytic (Kohut, 2009) and cognitive and behavioural traditions (Young,

Klosko, & Weishaar, 2003), that, for example, emphasise the role of the self in interpreting personal experience and shaping behaviour.

One final implication of the disease model is that it shapes expectations about the respective roles of patients and clinicians (Agich, 1997; Sadler & Agich, 1995). According to the disease model, individuals who experience mental illness are suffering a physical disease like any other and so are largely passive recipients of treatment. They therefore have relatively limited responsibility for their condition, beyond seeking and complying with empirically supported treatment. In summary, although the disease model tenets appear esoteric, in fact they profoundly influence understandings of what mental illness and recovery is, the legitimate objects and methods of mental health research, and the respective roles of professionals and patients.

Criticisms of the Disease Model

Criticisms of some or all of the tenets of the disease model are varied and many. These critiques are not merely claims that the disease model tenets are basically correct, but not yet realised, or inappropriately operationalised in current practice. Instead, these criticisms are that one or more of the tenets are fundamentally flawed assumptions and will never provide a coherent model for mental illness, irrespective of neuro-biological advances or nosological refinement.

Socio-political critiques.

Socio-political critiques typically reject internalism. They argue that mental illness is not real—not in the sense that mental suffering is not real—but that medicalising psychological suffering wrongly locates the cause of suffering within the mind. Some claim that “mental illness” is more accurately understood as a “life problem” (Szasz, 1961), an outcome of labelling processes (Goffman, 1963; Scheff, 1974), or unresolved family conflict (Laing, Esterson, & Cooper, 1970). Others point to more specific institutional systems that

exploit the disease model as a rhetorical device for financial gain (Conrad, 2005). For example, many have noted the extraordinary power of pharmaceutical companies with vast financial incentives to develop cures in search of a disease and so distorts clinical research and practise (Frances, 2013). Other critics highlight the undue influence of the insurance industry in influencing the creation of DSM (Kirk, Gomory, & Cohen, 2013).

It is well beyond the scope of this thesis to address all of these (what I gloss as) “socio-political” critiques of the disease model. Briefly, some commentators note these critiques can trivialise the reality of mental suffering (Pilgrim, 2013), provide little or no positive agenda for sensitively responding to mental suffering, and can overstate the explanatory power of socio-political factors (see Fulford & Sadler, 2000 for a review). Historically, there *are* egregious instances of politically contrived mental illnesses, such as drapetomania—a condition once thought by confederate doctors to cause slaves to run away from their masters (Kovács, 1989). But it is much harder to show how such socio-political processes can readily explain the range and specificity of symptoms observed in other, diverse mental conditions. A socio-political account explains avoidance in slaves, but founders as a precise account for the tics of Tourette’s syndrome or the word salad of florid psychosis. Syndromes of psychological disturbance in various forms have existed long before the advent of modern psychiatry, the pharmaceutical and insurance industries (Harris, 2013; Nasser, 1987). Similarly powerful social and political forces certainly shaped the occurrence and features of these syndromes in the past (Cohen, 1993), but acknowledging this does not amount to explaining the complex patterning of psychological signs and symptoms observed historically, and cross-culturally (Steel et al., 2014). These externalist critiques have therefore largely failed to threaten the ascendancy of the disease model.

Engel's bio-psycho-social critique.

A less radical criticism of the disease model is that it neglects the psychological and social aspects of illness experience. This was the impetus for the bio-psycho-social model developed by Engel (1980). Engel's model rejects the tenet of essentialism on the grounds it provides only a partial scientific account of disease and, as a practical consequence, contributes to a widely perceived disconnect between human experience and scientific, medical practice. Engel went on to argue that human health needs to be understood within a systems perspective (von Bertalanffy, 1956)—in which a biological understanding of disease is integrated into an understanding of the psychology of the person (experience and behaviours) and social context (including the immediate interpersonal environment, family, community and wider institutional influences). Notably, each level of the system is nested within higher levels of the system, and elements at different levels are thought to influence lower levels.

The main criticism of the bio-psycho-social model is that, in fact, it is not a “model” at all but rather a vacuous exhortation that everything is important (McLaren, 2006). Since it provides no guidance as to how, why or when different levels can or should be integrated, there are likely to be as many models as there are users of it—a situation that “borders on anarchy” (Ghaemi, 2009, p. 3). Ghaemi further argues that by permitting theoretical eclecticism it, ironically, perpetuates the very dogmatism and lack of integration Engel seeks to discourage. Indeed, some attribute the model's initial success and influence not to its substantive import, but rather its utility in maintaining an uneasy rapprochement within the discipline of psychiatry—between those with biological (“bio”) and psycho-analytic (“psycho-social”) sensibilities (Pilgrim, 2002). This leads to the further line of criticism that in practise the “psycho” and “social” elements are invariably collapsed as a singular level, the “psycho-social”, and added as a placatory afterthought to analyses conducted at the “real”,

biological level (Pilgrim, 2002). While this is clearly contrary to Engel's intent, it is consistent with Ghaemi's claim that the lack of specificity makes it an easy "model" to endorse regardless of one's practise or theoretical commitments.

Empirical critique.

A more serious challenge to the disease model is the increasingly influential view that an essentialist model of natural kinds provides an empirically flawed account of mental illness (Beebe & Sabbarton-Leary, 2010; Godman, 2013; Haslam, 2002, 2014; Kendler et al., 2011; Zachar, 2000), and is a metaphysically problematic notion more generally (Hacking, 2007). The empirical challenge to essentialism was foreshadowed as early as 1912, by the German psychiatrist Alfred Hoche, in his paper "The significance of symptom complexes in psychiatry", while commenting on the discovery that a particular kind of dementia—progressive paralysis—is caused by syphilis:

The main example of a happy final definition of disease conditions, which in all directions constantly prove to belong together, has been progressive paralysis. The success achieved here has perhaps been a misfortune in its side effects because it nourished the illusion that something similar might soon be repeated. (cited in Sass, 2009, p. 139)

More than a century later, Hoche's misgivings about the prospects for essentialist accounts of mental disorders have not abated. No more psychopathogenic spirochetes or infectious agents have been found (Zachar, 2000) and there are still no clinically reliable, biological tests for any psychiatric disorder (Lakhan, Vieira, & Hamlat, 2010). It was hoped that the advent of modern genomics would provide the first opportunity to empirically test the thesis of bio-genetic essentialism. If genetic essentialism is true, then specific gene main effects and gene x environment interaction effects should be associated with psychopathology. And so some speculated as recently as fifteen years ago that advances in genome

sequencing would mean that “in a few years, many areas of psychology will be awash in specific genes responsible for the widespread influence of genetics on behaviour” (Plomin & Crabbe, 2000, p. 806).

This optimism has been dashed. Genome-wide association studies (GWAS) provide no evidence that single genes of large effect can explain any mental illness (Ripke et al., 2013). Instead, the evidence shows that for major psychiatric disorders, as for any complex behaviour, at least hundreds or thousands (Chabris et al., 2013) of genes, each of very small effect, are involved. For example, Munafò, Zammit, and Flint (2014) note that, on average, particular gene variants explain less than 0.5% of phenotypic variance in mental disorder. Nor is this just an artefact of the limitations of how the phenotype is characterised. Recent GWAS linking genetic polymorphisms to 17 extensively studied “endophenotypes” (biological markers that might serve as objective markers for psychopathology) find a similar pattern: endophenotypes are “massively polygenic” (Iacono, Vaidyanathan, Vrieze, & Malone, 2014, p. 1346). These findings alone are a severe blow for attempts to biogenetically define mental illness.

But the full scale of empirical challenges for biogenetic essentialism becomes clearer when one considers that most genotype-phenotype main effects have not been replicated (Ioannidis, 2007) because even large studies, involving tens of thousands of participants, are typically underpowered. The problems are made vastly more complex again when gene-gene and gene-environment interactions are taken into account, requiring sample sizes of millions of participants to address even moderately complex questions (Burton et al., 2009; Chanock et al., 2007; Hein, Beckmann, & Chang-Claude, 2008). More troubling still, to robustly control for the confounding of genetic main effects by gene-environment correlations requires comparisons of gene associations in twins reared together and apart. Obtaining massive sample sizes for GWAS is hard enough, but attempting to do so for twin studies is

manifestly unfeasible (Turkheimer, 2011). In sum, there is no evidence that necessary and sufficient bio-genetic essences can explain mental illness, and the prospects this conclusion will be revised in the foreseeable future are dubious at best (Joseph, 2012).

These GWAS findings are broadly consistent with recent taxometric analyses. As mentioned earlier in this chapter, the existence of a latent taxonic structure is a necessary but not sufficient condition for the existence of some discrete essence (Haslam, 2014). The most up-to-date summary of this evidence from 177 studies involving more than 500,000 participants, suggests that the prevalence of taxonic disorders is about 14 %, and can be reliably observed in schizotypy, autism and substance abuse disorder. Again, this does not imply that these disorders necessarily emanate from discrete aetiologies, but rather shows how most clinical psychological disorders as we know them are certainly multi-factorial. In sum, bio-genetic and taxometric evidence provides convergent evidence that essentialist accounts of mental illness are empirically untenable (Haslam, 2014).

Homeostatic Property Clusters—an alternative to essentialist natural kinds?

Boyd's Homeostatic Property Cluster account.

Given these problems, an increasingly popular alternative to essentialism within the discipline is the so called homeostatic property cluster (HPC) account of natural kinds (Boyd, 1991, 1999). The account was developed to provide a non-essentialist explanation of the entities observed in complex systems. The term HPC stems from the observation that parts of complex systems can be more or less stable, not because of an underlying essence, but because of a network of homeostatic mechanisms. Hence, stability in mental illness might be explained in a similar way to how body temperature is maintained at a set point: not by a localised temperature regulator, but by a series of mechanisms that are triggered by negative feedback loops (vaso-regulation, perspiration, changing clothes, building shelter).

More formally, Boyd defines HPC natural kinds according to 10 principles. For present purposes, these can be summarised by the following claims that an HPC natural kind exists if:

1. a property cluster exists: i.e., a set of properties commonly co-occur;
2. one or more similarity generating mechanisms exist: there is one or more causal structure/mechanism to explain why the properties in (1) co-occur, and;
3. the cluster is causally important: i.e., it (probabilistically) explains, predicts or controls phenomena that are relevant to a particular discipline (the accommodation thesis).

The glue that holds the properties described in claim 1 together is not a singular essence, as in the essentialist account, but rather the homeostatic processes that explain similarities (claim 2). The account is permissive in the sense that not all of the properties need to occur in every member of the kind: there may be multiple mechanisms that cause the properties to cluster together (not all of which have to be present), and some of the properties may themselves cause the co-occurrence of other properties. For example, sleeping difficulties in depression may not just be one of the properties within the cluster that characterises major depression, it may itself be one of the mechanisms that causes other depressive symptoms (e.g. psychomotor retardation) (Wichers, 2014). The definition of a particular HPC, and the decision about which properties and causal mechanisms are relevant, is therefore an empirical question rather than a prior assumption. It is also non-reductive since properties within the cluster may be a part of the causal mechanism that fulfil requirement 2. The conceptual contrast between the causal processes that underlie HPC kinds and the essentialist natural kind account of the disease model are shown in Figure 2.

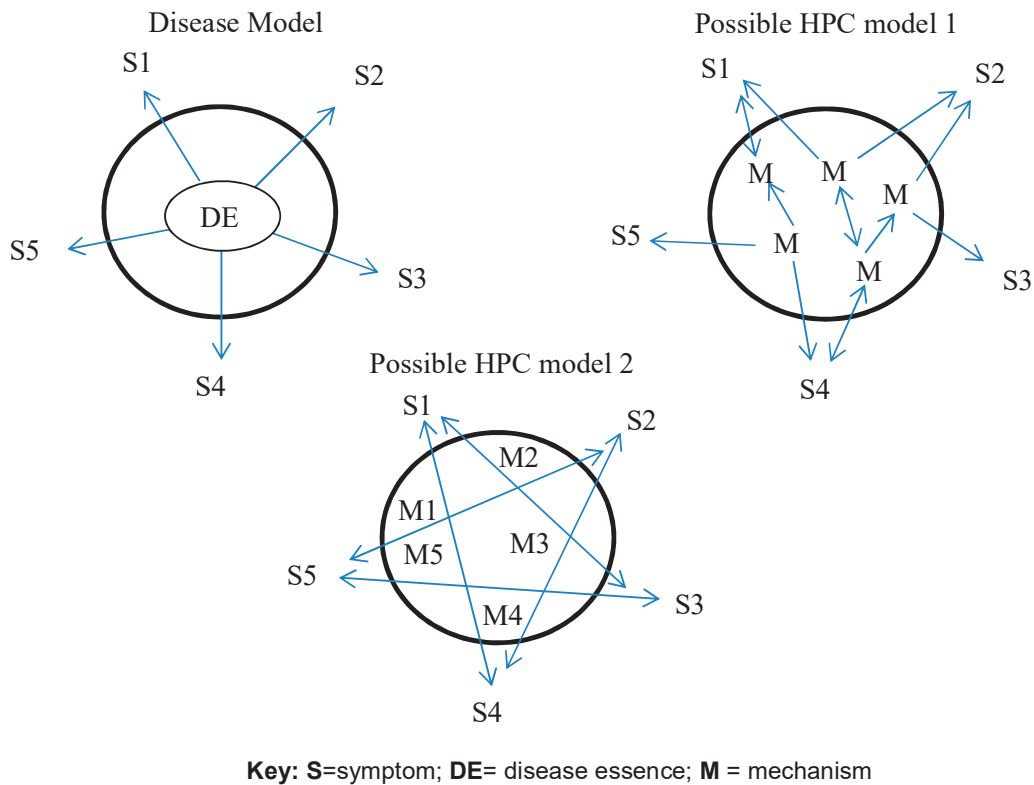


Figure 2: The disease model and two possible HPC models.

The standard Disease Model diagram shows how the disease essence is the causal antecedent of all of the symptoms 1 to 5. By contrast, the possible HPC model 1 depicts a series of mechanisms (M1-5) which jointly cause symptoms 1-5, and also interact with each other. Possible HPC model 2 shows how symptoms themselves may serve as mechanisms that cause other symptoms, and so mutually reinforce the symptom clustering (e.g. rumination may cause sleep disturbance, which causes low mood, which causes further rumination).

Accommodation thesis. The third key element to the HPC account is the so-called accommodation thesis, that is, “how schemes of classification contribute to the formulation and identification of projectible hypotheses” (Boyd, 1999, p. 147). The implications of the accommodation thesis are quite radical. Specifically, whether or not something is an HPC kind is partly the question of whether, within a particular discipline, that kind helps to explain, predict or control some phenomenon of interest to the discipline. On this account,

lilies and onions form a kind for botanists, but not for food scientists; “race” may be a kind for sociologists and historians, but not for biological anthropologists and geneticists. This adds an importantly pragmatic and perspectival element to kinds, that, as many have noted (Craver, 2009), is at odds with the traditional essentialist hope that kinds can be defined independently of human conventions or purposes. Instead, the HPC account navigates a middle ground between nominalism and essentialism. HPC natural kinds are not purely a matter of convention (as in nominalism) because they require there to be some empirically discoverable mechanisms that drive the clustering of properties.

Moreover, because HPC kinds, by definition, track causal mechanisms, they underwrite their claim to non-arbitrariness because they are (probabilistically) predictable and predictive. If, instead, the clustering of properties was co-incidental then it would not count as a kind. In this sense, the account is realist. However, by embedding the ontology of HPC kinds in human, epistemic concerns (i.e. concerns about whether a kind is predictive), it invites the question: predictive for whom? The account provides no basis for privileging a particular answer to this question and so inevitably leads to a disciplinary pluralism, constrained only by the predictable and predictive requirement (Craver, 2009). HPC natural kinds are therefore inextricably tied to both the world and human conceptions of it.

Boyd’s HPC account is strikingly similar to various frameworks within cognitive psychology to explain the process of categorization more generally (Barsalou, 1990; Bruner, 1957; McGarty, 1999; Turner, et al., 1987). McGarty describes these various cognitive accounts as constraint relations models. As with HPC accounts, these models were developed in part to address the inadequacy of essentialist accounts of categorization. As shown in Figure 3, the HPC constructs of clustering, causal mechanisms, and accommodation, parallel the three constructs that constrain and inform categorization within the constraint relations framework—perceived equivalence, background theoretical

knowledge and category use. According to constraint relations models, the perceived similarity between the observable properties of an entity parallel the requirement within HPC models that a set of properties should “cluster”—in the sense that similar properties should tend to co-occur across members of a category. According to the HPC account, there should also be a series of causal relationships that *explain* the clustering—elaborating how and why the set of properties tend to co-occur. This parallels the role of background knowledge within constraint relations models, which take the form of various beliefs about category members in a given context, often in the form of causal statements that imply the observed patterns of covariation. Finally, the accommodation thesis requires that the category must be *meaningful* in the sense that it must predict, explain, and control phenomena of interest to a discipline. In a similar vein, the constraint relations framework suggests that categories tend to be formed and used to the extent that they are explanatorily powerful to a subject in a given context. If a categorization scheme fails to be useful then this provides new background knowledge suggesting the causal theories that informed the categorization scheme to begin with are likely to be false. I will later elaborate the significance of these parallels, but for now just note that HPCs can be understood as a constraint relations model for explaining category formation and use within scientific disciplines.

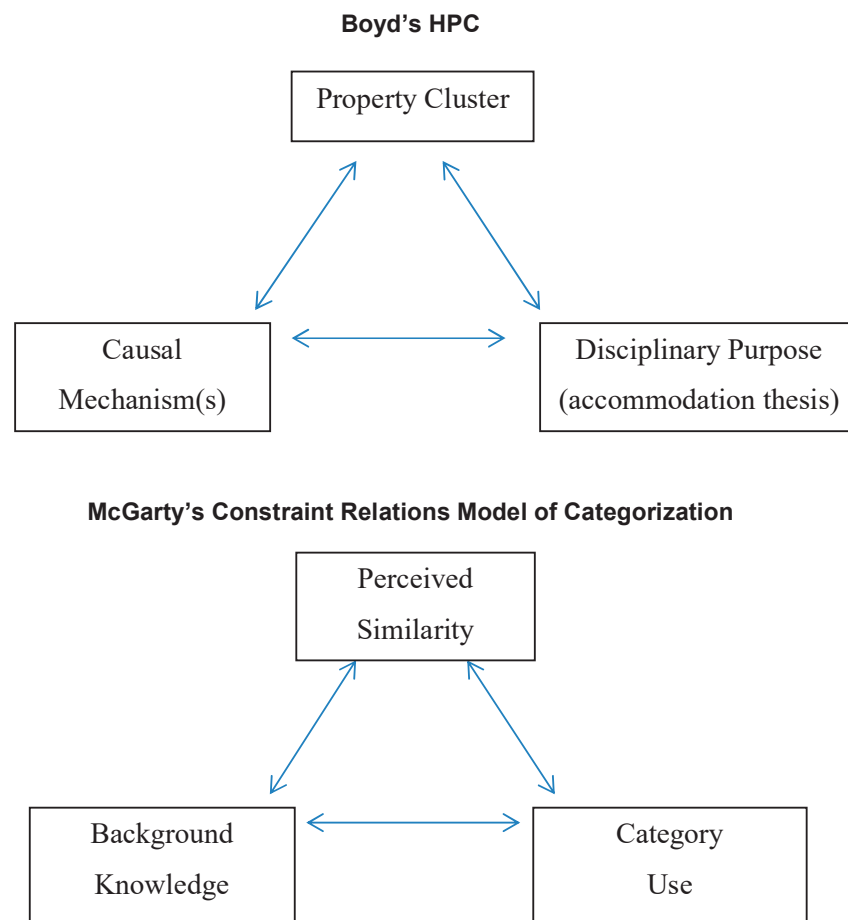


Figure 3: Parallels between HPCs (Boyd, 1991) and constraint relation models of categorization (McGarty, 1999)

One of the profound advantages of the categorization-based HPC account is that it allows for the sorts of more or less stable kinds (mental illnesses, biological species, rainforests) that lack necessary and sufficient essences but are central constructs within so many studies of complex systems. Scientists in these disciplines are assured that their subject matter is real and has a reasonably respectable metaphysical grounding. The cost, of course, is that it does not provide objective criteria for defining each kind's extension. Since multiple causal mechanisms can generate the clustering, none of which are necessary for generating the kind, one cannot be sure, for example, that even some non-psychological disease

categories carve nature cleanly at its joints (Williams, 2011), nor indeed that Plato's analogy continues to be apt. If natural kinds can be generated by multiple mechanisms, then such kinds may yield myriad latent forms on the taxon–continuum spectrum: as fuzzy kinds or, perhaps less commonly, discontinuous kinds. So long as there are multiple mechanisms involved in the genesis of the kind, there are likely to be borderline cases that cannot, in principle, be clearly defined as category members or non-members (Haslam, 2009).

Curiously, Haslam (2014) seems to suggest that HPC natural kinds must be at least “fuzzy kinds”. However, this seems to preclude the possibility that HPC kinds could also be pragmatic kinds (Zachar, 2000): kinds that derive from practically useful cut-points on an underlying, smooth continuum. If hypertensive signs and symptoms tend to co-occur, are causally generated by various mechanisms, and some particular cut point (140 on 90, say) is pragmatically useful for predicting clinical sequelae (heart disease, stroke) and in a way that optimises trade-offs between costs and benefits of interventions, then hypertension seems to qualify as an HPC natural kind. Insisting that the kind has a fuzzy or taxonic latent population structure seems to be an overly restrictive reading of Boyd's HPC account. Boyd's definition of cluster makes no reference to quasi-taxonicity as a necessary minimum condition for the presence of an HPC kind. Moreover, the idea of practical kinds is fully consistent with Boyd's accommodation thesis.

These points also reflect the inherent vagueness of the boundaries of “a mechanism” (Craver, 2013). Mechanisms themselves are not essentialist natural kinds, they invariably bleed into other mechanisms, and the magnitude and direction of their effects often depends on the magnitude and stability of various background factors. It is conceivable, then, that the same mechanism could generate either heterogeneity or homogeneity, depending on the context. In a sense, this is no problem for the HPC view, so long as the putative mechanisms are stable enough to generate probabilistically predictable and theoretically relevant causal

effects within a discipline then the kind continues to exist. If this means that the existence of a mental illness is spatio-temporally circumscribed, and context dependent, then so be it.

In sum, HPC natural kinds provide a plausible, non-essentialist ontology for studying complex systems like human psychopathology. Crucially, the HPC account preserves the realist intuition that mental illnesses are not mere human conventions because they must track mechanisms in the world. On this basis, Craver (2009) notes that “something like the HPC view is the most promising option available at present (for explaining mental illness)” (p. 591), and the approach, or accounts like it, such as Borsboom and Cramer’s (2013) causal network approach, is increasingly endorsed as a coherent alternative account of most if not all types of mental illness (Beebe & Sabbarton-Leary, 2010).

However, if a mental illness is best accounted as an HPC, the implications for many of the commonly, if tacitly, held disease model tenets, are potentially profound. It overturns not just essentialism but also discontinuism, and ahistoricism while also permitting non-reductive and externalist accounts of mental illness. Moreover, as Craver (2009) warns, the HPC account is unlikely to resolve nosological disputes, such as whether DSM is correct, because there will be “multiple incompatible answers to these questions depending on which mechanism one attends to, on how one describes the phenomenon, and on where one draws the boundaries of the mechanism” (p. 591). Answers to all of these questions depend on the disciplinary lens one adopts, and the level of consensus within a discipline about what needs explaining.

Kendler et al.’s HPC account of mental illness.

This ambiguity is highlighted in an attempt by Kendler et al. (2011) to apply an HPC-like account to the case of mental disorder. To understand their account, it is important to consider Kendler’s previous discussions on the limits of reductionism. Kendler argues that failures to find single genes of large effect to explain mental illness are humbling and that

researchers should resist calls to make psychiatry more vigorously biogenetically-based (as proposed by Kandel, 1998, see Kendler, 2005b) since more of the same will likely yield more of the same. Instead, Kendler (2005b) claims to be a non-reductive materialist who advocates a seemingly sensible third way: rejecting the dichotomies of either biogenetic reductionism or the anti-empirical excesses of radical social constructionism. Such a third way embraces the inherent and deep complexity of psycho-pathology, acknowledges that big, simple explanations are likely to be wrong and that progress will only be made if researchers aim for patchy reductions and piecemeal integration within an overarching, pluralistic and multi-level explanatory framework. Kendler notes as an example that if the discipline of psychiatric research operates at 15 levels (1 being the level of DNA and 15 the level of individual clinical manifestation) then rather than try to explain level 15 using level 1, researchers need to explain bits of this chain—one researcher examining levels 1–3, say, by explaining gene transcription; and another researcher examining levels 13–15, for example, by examining neuropsychological deficits in schizophrenia. Importantly, neither level has a privileged explanatory role. Kendler can therefore be characterised as rejecting essentialism and ruthless methodological reductionism, while maintaining his commitment to internalism.

To conceptually ground an alternative to essentialism, Kendler et al. (2011) propose that mental disorders are best understood by what they call “mechanistic property clusters”—explicitly inspired by Boyd’s HPC account. Kendler et al. go on to claim that psychiatric disorders “are objectively grounded features of the causal structures of the mind/brain” (p. 1147). However Kendler et al. seem to assume that HPC accounts are uncontroversially compatible with internalism, and so mental disorders should straightforwardly be construed as causal structures of the mind/brain (by which I take it they mean individual mind/brains—consistent with Kendler (2005b)). Kendler et al. go on to note the “potential relevance” of different kinds of causes, including the psychological and the social, but argue that these have

“underlying etiologic pathways (genetic, physiological and cognitive-affective)” (p. 1148). The use of the term “underlying” again suggests internalism, but is a puzzling choice in the context of an HPC-like account of mental illness—since, as I have argued, there is nothing necessarily “underlying” about HPCs. Limiting their analysis to causal structures within the mind/brain would be more clearly understandable if they acknowledged their interest was in addressing issues within molecular-genetics, neuro-psychology, or perhaps cognitive psychology. But they do not. Instead, they purport to offer a general account of what psychiatric disorders are and how they should be understood. It would seem, then, that Kendler et al. attempt to replace the static essences of the classic disease model with HPCs, but do not consider the implications for their internalist commitments.

Criticisms of HPC internalism

What I will call “HPC internalism”, developed by Kendler et al., is more deeply problematic than I have already sketched, for at least three reasons: internalism is arbitrary, will likely be empirically limited in general and will specifically fail to account for collective intentionality. I will elaborate these criticisms in some detail, in part to show the shortcomings of Kendler’s attempt to provide a non-essentialist account of mental illness. However, my criticisms also play a generative role by helping to positively define and justify the alternative account of mental illness I develop in the final section of this chapter. In particular, I hope to show why abandoning internalism as a meta-theory will enable a more explanatorily powerful science of mental illness. Moreover, I aim to cut through superficial agreement about the importance of social and environmental factors by diagnosing why this seeming agreement occurs, locating where substantive debate arises, and explaining why taking the social seriously (and so also causal processes outside the mind) is a bigger (and more useful) commitment than some suppose.

Arbitrariness objection to HPC internalism.

The first problem with Kendler's HPC internalism is that treating the boundaries of individual mind-brains as special is unnecessarily arbitrary. This is not an empirical claim, but rather a meta-theoretical point about delimiting the scope of scientific enquiry *a priori*. Kendler argues that the failure of reductionist accounts of disorder suggests the scope of enquiry should be broadened to consider macro-level phenomena on an equal footing to more micro-level phenomena. In Kendler's terms one should not privilege lower levels (Level 1, the level of DNA) over higher levels (Level 15, individual clinical manifestation). Rather, the relevant criteria for incorporating a particular level of analysis should be its *explanatory power*. But Kendler does not explain why Level's 1-15 are special and why, by implication, inter-individual and intergroup phenomena (Levels 16-17) should not also be considered within the scope of enquiry if these levels provide additional explanatory power. An alternative, less arbitrary, starting point for analysing psychiatric illness would be open to including the homeostatic causal mechanisms that operate *between* (groups of) individuals, in addition to the intra-individual processes. Only subsequent empirical research would clarify whether and to what extent inter-individual and inter-group processes play relevant causal roles for a specific condition.

A response to the arbitrariness challenge would be to insist individual human mind/brains are so causally complex they should be at the heart of scientific accounts of psychiatric disorders. But there is no prior reason to believe that rich, causal complexity suddenly stops, or is of more limited import, once the boundary of the skull (Kendler's level 15) is reached. In fact, as I will argue shortly there is much empirical evidence to the contrary (see Butler, 2011, for a review). At the very least, then, an internalist HPC account of psychopathology needs to *justify* the assumption that we should expect to find relevant causal processes only within individual mind/brains.

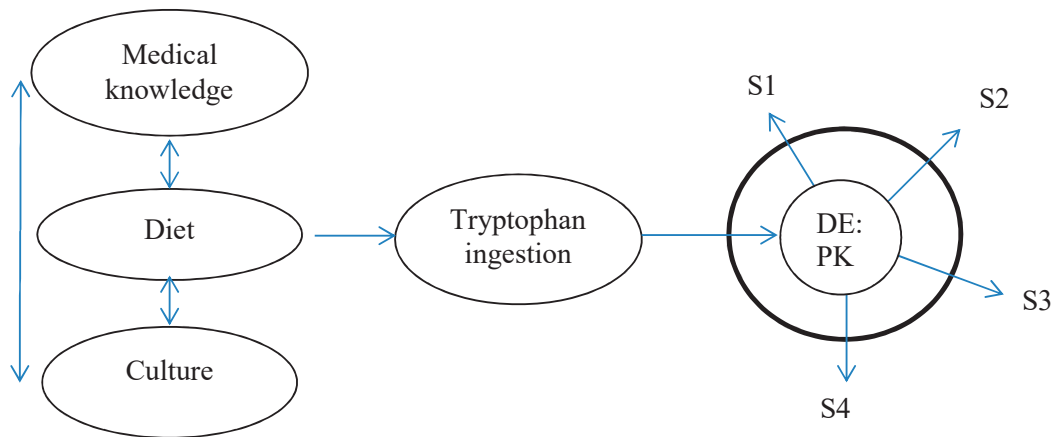
Empirical inadequacy objection to HPC internalism.

Privileging causal mechanisms within individual mind/brains is also likely to be inadequate empirically. While the assumption of internalism is, arguably, justifiable within the framework of an essentialist account of mental illness, it will likely lead to incomplete empirical explanations within an HPC framework. But before elaborating this argument more fully, it is important to neutralise an at times confused debate about the role of the social.

All sophisticated accounts of mental illness assume the social is important. Even those who Kendler criticises for advocating a hard biomedical reductionism, such as Kandel (1998), insist that “regulation of gene expression by social factors makes all bodily functions, including all functions of the brain, susceptible to social influences” (p. 461). Similarly, Kendler (2005a) himself has persistently argued that social and political processes are critically important and causally influence mental illness. At first glance, it seems hard to reconcile this with charges of reductionism or internalism. How is it even possible to admit a causal role for the social, while insisting on internalism? I will explain that this is possible and coherent, but only if essentialism is assumed.

There can be no doubt that social and environmental influences can ultimately trigger gene expression. Certain mutations of chromosome 12, for example, can cause phenylketonuria (PKU)—a neurological disease that leads to intellectual disability and seizures—but only if an individual’s diet contains high levels of the amino acid tryptophan. But this does not imply that the social has an *intrinsic* causal role. As shown in Figure 4, a complex chain of social and historical factors may determine whether the knowledge and medical testing is available to diagnose PKU, whether individuals have resources to access diagnosis and the means to limit tryptophan in the diet. But the proximal causal processes are

still highly localised: the presence of the PKU mutation and tryptophan is a necessary and sufficient condition for developing Phenylketonuria symptoms.



Key: DE = disease essence; S = symptom; PKU = genetic mutation in 12q22-q24.1

Figure 4: A simplified hypothetical model of the social and cultural determinants of the genetically determined disease Phenylketonuria.

The external determinants of the triggering conditions can be partitioned off from explanations of the phenomenon itself, while at the same time insisting that *of course* they are profoundly important in explaining particular symptomatic cases. It is this sort of reasoning that permits an avowed internalist and reductionist to acknowledge the rich complexity and importance of social and environmental causes, while insisting that these influences are ultimately mediated by the disease-causing essence. This gene-centric view of the role of social and environmental causes is summarised well by Kandel: “all of ‘nurture’ is ultimately expressed as ‘nature’” (1998, p. 457).

But Kandel’s (1998) argument seems plausible in the context of mental illness only because it presupposes the existence of essentialist natural kinds. Remember, essentialist natural kinds are inductively potent precisely because they permit law-like predictions given precise initial conditions (gold melting at 1064 degrees centigrade). Regardless of whatever

complex, distal mechanisms influence the presence or absence of these initial conditions, the disease phenomenon can be localised and proximal antecedents and consequences of the thing-in-itself can be plucked out. However, if, as the HPC account of mental illness requires, talk of necessary and sufficient essential conditions is abandoned, then it is much harder to neatly partition off localised processes within the brain as being somehow more fundamental or intrinsic to the phenomenon.

Consider, for example, the HPC account of depression outlined by Wichers (2014). Here, the causal processes that help define the phenomenon are spatio-temporally dispersed and unfold in intricate ways over time. Sleep disturbance may lead to difficulties concentrating, leading to feelings of humiliation, which in turn triggers a stress response, and further sleep disturbance. According to the HPC account, none of these links is necessary nor sufficient to define depression and there is no single point in the chain where the depression “really is”.

But as a result, and quite unlike the case of essentialist natural kinds, the distinction between proximal and distal, and antecedent and consequence, are fuzzy. Intricate feedback loops, over varying spatio-temporal scales, just *are* the thing-in-itself, not a background enabling condition for it. Moreover, these feedback loops are inextricably embedded in external social interactions (see Butler, 2011 for a review). For example, evidence that the physiology of mothers and infants can become dynamically coupled has been shown on measures of stress reactivity (Waters, West, & Mendes, 2014) and cortisol levels (Hibel, Granger, Blair, & Cox, 2009; Williams et al., 2013). And similar effects have been shown for romantic couples in relation to heart rate (Ferrer & Helm, 2013; Helm, Sbarra, & Ferrer, 2012), respiration (Ferrer & Helm, 2013), voice frequency (Gregory & Webster, 1994), and self-reported affect (Butner, Diamond, & Hicks, 2007). In sum, HPC mechanisms are spatio-

temporally diffuse, rather than localised, and their boundaries cannot be delimited by appeals to proximity (see also Craver, 2009).

An HPC account of mental illness, then, is not a metaphysical trick for glossing the lack of evidence for an essentialist account of mental illness. Instead, the account requires that the notion that mental illness is a complex system be taken seriously, and that conceptual and methodological strategies be revised accordingly. Bechtel (2009) neatly summarises the explanatory power to be gained by this approach. He argues that recent interest in mechanistic accounts of psychological processes might suggest a privileging of “looking down” (reductionism), but the empirical reality is that it necessarily involves the complementary processes of “looking around” (reconstructing components of the mechanism and examining their modulation by external/environmental inputs) and “looking up” (situating the mechanism in context). Instead of Kendler’s “patchy reductions” and “piecemeal integrations” (Kendler, 2005b, p. 15), Bechtel argues that all three of these distinct methodological strategies, in concert, are needed to develop a rich explanatory account. Elsewhere, Bechtel examines numerous examples from cognitive neuropsychology (see also Frisch, 2014) to illustrate the empirical generativity of, in particular, situating mechanisms in context, to adequately explain circadian rhythms (Bechtel, 2009, 2013; Bechtel & Abrahamsen, 2010), visual perception and memory consolidation (Bechtel & Abrahamsen, 2008).

This leads again to the accommodation thesis within HPC: what is it that psychology and psychiatry are attempting to explain? By Kendler’s (2005a) own reckoning, the majority of variation that explains the presence or absence of major psychiatric conditions, including schizophrenia and major depression, is social and environmental. If one’s primary interest is in developing biochemical interventions, then this variation is noise to be partialled out. If instead, one’s interest is in social interventions, then social variation is signal to be analysed.

But either way, if one seeks to adequately explain, predict or control mental illness, one needs to take seriously the idea that social mechanisms, external to the brain and body, might be intrinsic to the phenomenon of interest. As I have argued, this is a much bigger commitment than the platitudinous—and implicitly essentialist—acknowledgement that the social matters because it can trigger underlying predispositions. The language of “underlying” leads one to suppose that one can look down from symptoms to discover where inside individual heads the mental illness mechanisms really are. But this neglects the complementary strategies of “looking around” and “looking up” from symptoms—an approach that has been so empirically generative in studies of similarly complex phenomena (Frisch, 2014).

Testing the “looping” hypothesis: social identity as an HPC causal mechanism

I have argued that invoking internalism in an HPC account of mental illness is arbitrary and is likely to have limited explanatory power. The examples I have discussed so far suggest that HPC accounts of mental illness need to consider that causal mechanisms may be embedded in the immediate social context. In this section I discuss a good example of the sort of externalist HPC mechanism that should be explored empirically, but that an internalist HPC account precludes from consideration.

The example comes from the philosopher Ian Hacking (1999) who argues that shared lay beliefs may dynamically interact with expert beliefs in a way that substantively changes the properties of mental disorders. Hacking argues that humans are uniquely reactive to the very act of being classified, in contrast to “indifferent kinds”—collections of things that are unable to react to their own classification (books, mud, spirochetes). Hacking’s point is more radical than the claim that, unlike labelling books, disorder labels merely affect how people act—as in certain sociological theories of labelling (Goffman, 1963; Link, Cullen, Struening, Shrout, & Dohrenwend, 1989). Hacking goes the further step of claiming that individuals who are labelled as mentally ill can become collectively aware of their disorder label, confer

with each other about the meaning of the disorder (through patient support groups and lobby groups) and that this process will systematically change people and the very properties of the disorder itself. In this way, mental illness is what he calls a “looping kind”. The first arc of this loop extends out from a community of experts (doctors, administrators, and researchers), who observe and define mental illness empirically, and apply illness labels to individual sufferers. The second arc of the loop extends from these sufferers, who discuss their shared disorder experience, and collectively reinforce, resist, or otherwise react to a disorder label and thereby change the pattern of symptoms observed by clinicians.

Hacking argues these looping effects seem to explain how disorders can rise and fall within particular socio-cultural niches. For instance, multiple personality disorder was scarcely observed in the first half of the twentieth century. However, following lobbying by a group of psychiatrists, by the 1970s the disorder was officially included in DSM and, in the wake of both public media and increasing expert interest in the phenomenon, the number of cases increased from around 200 to 20 000 within the space of ten years, and the average number of alters displayed by sufferers increased from about 7 to 15 over the same period (Piper & Merskey, 2004). Hacking is careful to avoid the question of whether and to what extent this is attributable to increased rates of accurate diagnosis, or to the power of suggestion and iatrogenic treatment effects. Instead, Hacking notes that:

In 1955 this was not a way to be a person, people did not experience themselves in this way, they did not interact with their friends, their families, their employers, their counsellors, in this way; but in 1985 this was a way to be a person, to experience oneself, to live in society. (p. 299)

Hacking’s emphasis is on the idea that this was a way of being that was only made possible once the disorder had become widely recognised. Hacking goes on to discuss examples of similar processes for other disorders and social categories, including autism,

obesity, homosexuality, genius, and suicide. In all cases he presumes that the occurrence of the behaviours reflects complex bio-genetic processes in interaction with the environment and in all likelihood at least aspects of these behaviours have existed across time. But Hacking's point is that only after these categories became objects of collective awareness, did they become ways that people could intelligibly see themselves and that others saw them. This collective awareness of a category was sufficient to induce people to start behaving in ways they would not have had they been unaware of the category. The looping effect therefore acts as a feedback mechanism that can be sufficient to amplify and even change the symptoms displayed by entire groups of sufferers.

Currently, Hacking's postulated looping effect is philosophical conjecture supported by suggestive, sociological anecdotes. If these looping effects could be demonstrated empirically, then the theoretical implications for understanding stability and change in mental disorder are potentially profound. Perhaps most importantly, such effects would provide a disorder-specific socio-genic mechanism. By contrast, almost all other postulated social causes are generic. Income poverty (Reiss, 2013), inequality (Murali & Oyebode, 2004), and inadequate housing (Evans, Wells, & Moch, 2003) have all been demonstrated to put people at risk of various mental disorders, but none of these social causes is clearly disorder specific. An empirically well elaborated looping mechanism would therefore greatly enrich theoretical explanations as to how distinctive symptom constellations can emerge and change within particular socio-cultural niches. The question then, is whether these looping effects are logically coherent and empirically plausible.

Looping effects are theoretically compatible with an HPC account. Assessing Hacking's looping argument requires careful attention to the conceptual role that social processes play in the disease model and the HPC model. Presently, I show how Hacking's thesis can be refuted if mental illnesses are essentialist natural kinds, but that looping effects

are quite compatible with an HPC account. To demonstrate this, a useful distinction can be drawn between what Searle (1995) calls “social facts”, whose truth value depends on shared beliefs, and “brute facts”, which are true independently of the values of the observer.

According to Searle, social facts are made by assigning function to brute facts. This is what allows us to experience an inter-subjectively meaningful world—full of game show hosts, weeds and watering cans, rather than fleshy bipeds, spiky plants and concave bits of metal.

Even claims about heart disease, says Searle (1995), are social facts. That atherosclerotic plaques in arteries tend to stop hearts from beating is a brute fact about biology. But the further claim that atherosclerosis is a *disease* is teleological—since it supposes shared beliefs about the value of life, health and minimising pain. Similarly, psychiatric syndromes are diseases only if they cause distress or social and occupational dysfunction. Even though such claims appear to be brute facts (because, for instance, there is an empirically demonstrated inverse relationship between depressive symptoms and job output) there is still an underlying value judgement, signalled by the word “dysfunction”, that these consequences are undesirable. As mentioned earlier in this chapter, despite attempts to provide naturalistic accounts of functioning (Wakefield, 2007), very few find these plausible (Brülde, 2007; Houts, 2001; Murphy & Woolfolk, 2000). These naturalistic accounts fundamentally fail to explain why groups of humans care about whether we live long, short, productive or happy lives, in a way that nature does not.

But social facts do not just imply shared values towards brute facts. Particular kinds of social facts can initiate complex social and institutional responses. When a legitimate medical practitioner declares that “this counts as heart disease”, this, when uttered in a modern western medical clinic, confers on the phenomenon a particular symbolic status that can be sufficient to have profound causal effects, including visits to the hospital, insurance claims, and time off work. Not just anybody can make such declarations and initiate these

causal effects. According to Searle, claims of this sort are particular kinds of social facts, institutional facts, since they rely on shared agreement that particular experts (doctors) are authorised to make such declarations.

This distinction between brute and social facts allows proponents of the disease model to argue that supposed “looping effects” are overblown. The disease model proponent can concede that social groups govern what counts as dysfunction, and that psychiatrists are sanctioned to operationalise these values to construct institutional facts about disorders, and that groups of people can collectively resist these claims. But so what? This does not change the brute fact that a person with, say, untreated *Treponema pallidum* spirochetes will eventually develop dementia. Affected individuals can be collectively aware of the category “syphilis sufferer”, and contest its social meaning all they like: *T. pallidum* does not care. Beliefs about syphilis are causally interesting for the fate of *T. pallidum* only if an infected individual suspects something is wrong, seeks medical attention and initiates the cascade of institutional responses discussed above. In a similar way, if all psychiatric disorders are essentialist natural kinds then, in theory, the social meaning of symptoms can be endlessly contested or reinterpreted, but the underlying disease essences are brutishly indifferent. This allows the disease model proponent to block the looping argument, and adds grist to the essentialist mill by showing that descriptive nosologies are broken.

Kendler et al. (2012) assume that similar reasoning applies to HPC accounts of mental disorder. They argue that communities of experts, informed by wider social values, can decide whether a potentially psychiatric phenomenon is a disease, but once this is agreed, then scientists and practitioners can get on with the empirical job of describing and controlling the biological mechanisms within individual brains that cause symptoms to co-occur. Kendler’s account therefore suggests these social processes are no different for HPC mental conditions than for any other essentially defined, biomedical condition.

But if mental disorders are HPCs as I have claimed, there are reasons for doubting that things are so simple. Suppose, first, that something like Kendler et al.'s (2012) account is correct: there is a set of brute biological mechanisms in individual heads, shared to varying degrees by category members, and these mechanisms cause symptoms and these in turn have (societally agreed) dysfunctional consequences. Suppose then, that researchers and clinicians reify this symptom cluster by declaring it as an institutionally validated disorder. Hacking claims this may induce individuals to behave and feel in ways they might not have without collective awareness of the disorder. It is at this point that the disease model proponent can insist this is irrelevant—what matters conceptually is the presence or absence of the disease essence.

But the HPC internalist cannot make this move because there is no clear reason why this “looping effect” is not just another causal mechanism. Perhaps if it could be demonstrated empirically that an individual displayed sufficient signs and symptoms for disorder, but displayed no other, known, causal mechanisms, then that instance might be declared factitious or malingering. But this leaves a spectrum of individuals who may have relevant biological mechanisms, to varying degrees, but tip into or out of institutionally defined dysfunction only when a certain collective awareness is available to them. The HPC account then, properly understood, seems to be compatible with the claim that looping effects can serve as relevant causal mechanisms.

Shared beliefs as HPC causal mechanisms. If looping effects are to be considered seriously, it needs to be shown not just that they are theoretically relevant and compatible with an HPC account, but they should also be empirically plausible. Specifically, since shared beliefs play such a crucial role in looping effects, it needs to be shown that such beliefs are sufficient to cause the sorts of symptoms that are of interest to clinicians.

There can be little doubt that causal mechanisms that generate symptoms involve beliefs, desires and intentions. A now vast empirical literature demonstrates that the experience of physical and psychological symptoms is profoundly influenced by the beliefs, assumptions, and attributions of the perceiver (see Petersen, van den Berg, Janssens, & Van den Bergh, 2011 for a review). There is also extensive, high grade evidence that psychological therapies, which attempt to change individual beliefs, predict long term improvements in self-reported and clinician-rated outcomes for many psychiatric disorders (Cuijpers et al., 2013; Cuijpers, van Straten, & Warmerdam, 2007; Mitte, 2005). A growing literature further shows how these changes in beliefs and symptomatology are mediated by long term, structural brain changes (Buhle et al., 2014; Doehrmann et al., 2013; Hahn et al., 2015).

Although it might be granted that beliefs influence symptoms, it could be argued that these are not collectively shared, but are individualistically held, or at least that beliefs relevant to maintaining mental illness are. Indeed many theoretical analyses of symptom perception propose that illness representations are idiosyncratic memory traces or representational schemata (Brown, 2004). Similarly, Cognitive Therapy is predicated on the assumption that mental disorders are caused by individual cognitive distortions (Beck, 1995) or irrational schemas (Young, 1994) that need to be changed in order for individuals to recover.

However, an extensive literature demonstrates that socially shared beliefs have a causal role in emotional experience (Barrett et al., 2006). Classic experimental research by Schachter and Singer (1962), for example, found that when experimental participants were injected with epinephrine (a drug that causes physiological arousal), and provided no explanation for the possible symptoms they might experience, they were subsequently influenced by a confederate (who had allegedly received the same injection) displaying either

anger or euphoria. Participants' private ratings of their own emotional state, as well as observer ratings of participants' behaviour (blind to condition), aligned with the confederate's expressed emotion, but only when participants had not been told by the experimenter what side-effects they should expect. In other words, in the absence of a plausible explanation of their internal physiological state, participants looked to *similar others* to understand their inner experience, and this, in turn, changed how they felt and behaved.

Episodes of mass hysteria provide a dramatic illustration of just how quickly these shared beliefs can be inferred, transmitted and reinforced. These syndromes are typically characterized by the rapid onset of anxiety symptoms (especially dizziness, nausea, hyperventilation) that then become spontaneously "transmitted", line of sight, to adjacent others. These episodes have been very widely documented across time and cultures, including in groups of schoolchildren in the US (Jones et al., 2000), Africa (Ebrahim, 1968), Britain (Smith & Eastham, 1973), and Malaysia (Bartholomew & Wessely, 2002); among Australian airline workers (Balaratnasingam and Janca, 2006); in refugee camps in Nepal (Van Ommeren et al., 2001), and groups of nuns (Bartholomew & Wessely, 2002). Mass *motor* hysteria has also been widely observed, and is again socially transmitted, but displays more gradual onset and shared motor symptoms (shaking, twitching) that may persist over weeks and months (Wessely, 1987).

The socio-genic basis to these syndromes is clear: organic causes are ruled out, they spread when people see or hear about similar others being affected (Jones et al., 2000), and they tend to be transmitted from higher status individuals to lower status people within tight-knit communities (Bartholomew & Wessely, 2002). Some speculate these phenomena can be explained by individual personality factors (Wessely, 1987). But extensive reviews of the evidence for personality vulnerabilities in mass hysteria are weak and inconsistent (Ali-Gombe, Guthrie, & McDermott, 1996; Bartholomew & Wessely, 2002), leading

Bartholomew and Wessely to conclude that “there is no particular predisposition to mass socio-genic illness and it is a behavioural reaction that anyone can show in the right circumstances” (p. 304).

There is also evidence that these socio-genic processes have wide-ranging effects on more enduring, diagnosable mental illness. Recent research shows that, in a similar way, symptoms can spread through social networks and influence rates of depression (Rosenquist, Fowler, & Christakis, 2011), anxiety (Eisenberg, Golberstein, Whitlock, & Downs, 2013), and deliberate self-harm (Taiminen, Kallio-Soukainen, Nokso-Koivisto, Kaljonen, & Helenius, 1998). Although these correlational, network studies are in their infancy, a larger body of research investigating placebo and nocebo effects provides compelling experimental evidence that shared meaning plays a specific role in symptom experience in the context of mental illness (Barrett et al., 2006).

Many mental illnesses are highly susceptible to placebo effects, including schizophrenia (Reddy et al., 2011), chronic pain (Murray & Stoessl, 2013; Zhang, Robertson, Jones, Dieppe, & Doherty, 2008), major depression, and anxiety disorders (Kirsch, 2014). A large meta-analysis of randomised control trials, for example, suggests that approximately 80% of the treatment response to anti-depressants are placebo effects (Kirsch, 2014; Kirsch, Moore, Scoboria, & Nicholls, 2002) and this difference narrows to clinical insignificance when patients are given “active placebos” (placebos with mild side effects, but no known psycho-active ingredients) (Moncrieff, Wessely, & Hardy, 1996). Placebo injections provide more pain relief than placebo pills (Zhang et al., 2008); red, yellow and pink placebos tend to have a stimulating effect, whereas blue and green coloured placebos have a sedating effect (Craen, Roos, Vries, & Kleijnen, 1996); and capsules, branded tablets and larger tablets are perceived to be more pain relieving than smaller or unbranded placebos (Barrett et al., 2006; Branthwaite & Cooper, 1981). These differences in clinical outcomes are very difficult to

explain without appealing to varying social understandings about what these different placebo characteristics mean.

These examples illustrate the complex ways in which implicit, shared beliefs about how “we” behave and interpret bodily states can influence the perception, experience and transmission of emotional and somatic states. These findings provide convergent evidence for the claim that shared beliefs can play the sort of causal role required to generate looping effects. Of course, proponents of the disease model can dismiss these social processes as a superficial distraction from the underlying pathology. Again, this is because the disease model is not conceptually bound to symptomatology. But HPC accounts of mental disorder, dependant as they are on mechanisms that explain symptoms, cannot be so dismissive.

Mental disorders as shared self-categories. So far, it has been shown: that looping effects could provide an innovative explanation for stability and change observed in mental illness; that looping effects are theoretically compatible with an HPC account of mental illness; and finally that there is extensive empirical evidence showing how shared beliefs can play the sort of causal role in predicting symptoms, that looping effects require. But there is just one missing ingredient required for demonstrating the plausibility of looping effects. This is a more rigorous conceptual and empirical account of what it means to be “self-conscious” as a disorder sufferer. It is not enough to merely show that shared beliefs can influence the sorts of behaviours and emotions that typically count as symptoms. Hacking (1999) is clear that it is a particular kind of collective self-awareness as *a disorder sufferer* that is required:

I do not mean only the self-conscious reaction of a single individual to how she is classified. I mean the consequences of being so classified for the whole class of individuals and other people with whom they are intimately connected. (p. 115)

It is this collective awareness as a group of fellow sufferers that is required for looping effects to generate disorder-specific changes to the properties of a disorder category. However, as a number of commentators have noted, this sense of collective self-consciousness is conceptually vague (Kuorikoski & Pöyhönen, 2012). Hacking himself suggests that explaining this aspect of looping effects is a question for psychology and sociology—but provides few clues as to how this might practically be achieved. Perhaps for this reason, the looping argument has been relegated to philosophical speculation and has not received serious empirical attention.

In this thesis, I propose that Social Identity Theory (SIT) (Tajfel & Turner, 1979), and the related Self-Categorization Theory (SCT) (Turner et al., 1987), provide a theoretically well-developed means of empirically operationalising the kind of collective self-awareness that Hacking (1999) has in mind. This, so-called Social Identity Approach (SIA), provides an empirically well-supported framework for explaining how social categories become internalised as part of the self-concept and how, in turn, such internalisation enables shared beliefs, feelings and behaviour. Perhaps most importantly, the account does not merely assume that individuals passively conform to category stereotypes, as in labelling theories (Scheff, 1974). Rather, the account elaborates how and when category members can actively contest the social meaning of category stereotypes and in a way that systematically changes individuals' behaviour. SIA seems well placed, therefore, to empirically operationalise the kind of collective self-awareness required to test looping effects.

This crucial move is inspired in part by the observation noted earlier that the HPC account is a constraint relations model of categorization (McGarty, 1999). HPC disorder categories are developed by and for communities of expert academics and clinicians seeking to understand how observed regularities can explain, predict and control dysfunction (as defined by a wider society and operationalised by experts). But disorder categories can begin

to serve quite different purposes for other communities. I have touched briefly on how disorder categories are not only scientific constructs, they also serve a powerful, socially sanctioned, administrative purpose. Disorder categories allow diagnosed individuals, for instance, to access treatment and insurance, and be excused from work and social obligations (Parsons, 1951). In addition to this, Hacking's looping effects argument suggests that disorder categories can become subjectively meaningful from the perspective of the categorized—as self-categories. Being a depressive or running amok helps predict, control and explain to the self and others what it is that I do, think and feel.

But running amok or being depressed are not ways of understanding the self that just happen to arise within individuated heads, as if by coincidence. These categories explain and constrain what it is *we* do. There is nothing spooky about this empirically or metaphysically. Self-categories are trivially located within brains but are causally inter-connected by intricate, language-mediated signalling pathways (Postmes, 2003; Turner & Oakes, 1986). The looping effects argument implies that these extended causal mechanisms can systematically sustain and transform the properties of HPC disorders. In later chapters I will elaborate this point by showing how self-identifiers can use a shared disorder category to collectively negotiate and shape themselves and their social world in ways that extend and often profoundly challenge their original institutional functions: lobbying for access to particular kinds of treatment, supporting fellow sufferers to creatively reinterpret their condition, or even contesting social consensus about what counts as dysfunction—as when gay rights activists successfully lobbied to remove homosexuality as a DSM psychiatric disorder.

The SIA therefore provides a potentially powerful tool for testing looping effects. Theoretically, shared disorder self-categories are unique because they provide a means of understanding how processes outside the head can systematically shape HPC categories. But to date, the few theoretical models and empirical studies of HPC mental disorders assume

internalism, at least within psychology and psychiatry (Borsboom & Cramer, 2013; Cramer et al., 2010; Kendler et al., 2011). This arbitrary constraint does not even allow the looping hypothesis to be formulated. In this thesis, I therefore start to redress this limitation by providing a self-categorization account of Hacking's postulated "collective self-awareness" and do so in a way that unshackles HPC mechanisms from internalism.

Summary: towards a positive alternative HPC account

In this chapter I have explored a central problem in theoretical understandings of psycho-pathology. This is the challenge of explaining coherence, stability and change in mental disorders. In surveying this conceptual terrain, I have focused on two dimensions that map out important contours of the debate. These two dimensions, shown in Figure 5, relate to whether mental illnesses are best understood as essentialist natural kinds (essentialism–nominalism, on the x axis) and to what extent they are determined by processes within or beyond individual heads (internalism–externalism, on the y axis). This conceptual mapping is in some ways problematic. There are few unalloyed proponents of any of the positions I have discussed. Most all proclaim that the biological, psychological and social interact in complex ways. This is the anodyne bio-psycho-social consensus. Different meta-theoretical positions become perspicuous, however, when attention turns to assumptions about *how* these levels interact.

I began by outlining an important, and still highly influential, meta-theoretical stance: the disease model of mental illness. Its position in the bottom left quadrant of Figure 5, reflects two of the core tenets that underlie it—essentialism and internalism—and the related positions of methodological reductionism, discontinuism, and ahistoricism. By presupposing the existence of mental illness natural kinds, the disease model is well placed to explain coherence and stability in mental illness symptoms. Explaining cross-cultural and (sometimes rapid) longitudinal variation is more difficult. Proponents of the disease model

suppose this variation can be explained away by a combination of inadequacies in current descriptive diagnostic systems, reporting bias and because varying social and environmental influences trigger underlying disease essences. Despite the profound influence of the disease model, I have reviewed evidence showing that discontinuism and essentialism, in particular, are empirically untenable assumptions for understanding the vast majority of mental illnesses (Haslam, 2014). Specifically, as discussed, recent area-wide genetic analyses have confirmed that most mental illnesses, and endophenotypes, are massively polygenic. These empirical findings have contributed to what has been described by some as a crisis of confidence in diagnostic categories (Frances, 2013) and a view that conceptual understanding of mental illness is at a pre-Copernican stage (Pemberton & Wainwright, 2014).

In contrast to the disease model, I also mentioned various socially determinist, socio-political accounts, according to which mental “illness” is essentially caused by phenomena external to the mind/brain (shown in the upper left quadrant of Figure 5). Characterising this position as essentialist is perhaps controversial (although see Rangel & Keller, 2011). The point is simply that these positions claim mental illnesses are (to varying degrees) reducible to the effects of: illness labels and stereotypes (Scheff, 1974), or family conflict (Laing et al., 1970), or problems of living and political expediency (Szasz, 1961). “Mental illnesses” in these cases are normal reactions to particular abnormalities in the world. Among the challenges these socially determinist approaches face is in specifying falsifiable mechanisms for how these processes cause such distinctive symptom constellations.

Disorder Identity: Opportunity or Obstacle?

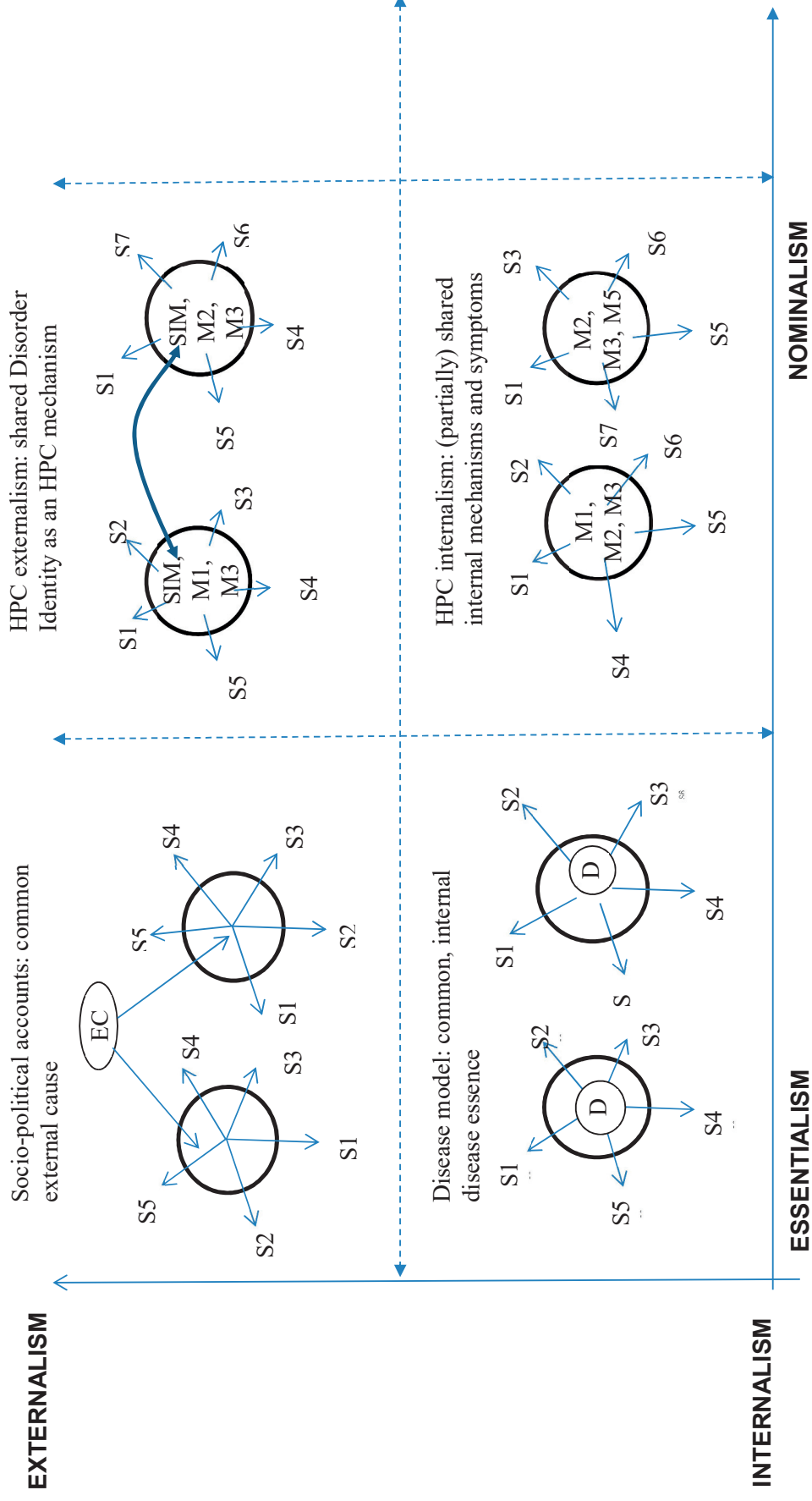


Figure 5: Causal models of mental illness: a two dimensional conceptual framework.

But giving up on essentialism in the wake of these challenges, particularly with the disease model, is not easy. Alternatives include radically relativist and nominalist positions, according to which mental illness kinds are arbitrary human constructs, unconstrained by an observer-independent reality. At their extreme, these approaches preclude empirical analysis of mental illness.

More moderate and empirically driven stances are possible. Boyd's (1991) HPC account is a particularly important theoretical contribution in this regard. The approach ties mental illness kinds to the process of categorization. As with categorization in general (McGarty, 1999), HPC kinds are constrained by: perceived equivalence, or clustering, of properties; by background knowledge about mechanisms that cause the clustering, and; by the purpose for which HPC kinds are used. This flips the standard essentialist account on its head. Rather than presupposing the existence of underlying essences that explain symptomatic regularities, the account starts with the regularities in symptoms and focuses the empirical task on discovering mechanisms that explain them—whatever form or forms these mechanisms take.

According to prominent recent attempts to apply such an HPC account to the analysis of mental illness by Kendler et al. (2011), this constellation of symptoms must be generated by causal mechanisms *within* the mind/brain. To show this, Kendler's account is located in the lower right-hand quadrant of Figure 5. I have taken some time to elaborate why Kendler et al.'s HPC account is problematic. While it is a positive advance to the extent that it provides an empirically grounded alternative to essentialism, and rejects ruthless reductionism, I argue it wrongly retains the disease model assumption of internalism. The debate is further confused by superficial agreement about the importance of social and environmental causes of mental illness. I have unpacked how this seeming agreement arises,

and why it belies a more substantive, and implicitly essentialist, assumption that the more causally potent forces must occur intra-individually.

I argue that HPC accounts need not assume an asymmetry between internal and external causes. Moreover, an internalist HPC account of mental illness cannot be sustained because it is arbitrary, likely to be empirically inadequate in general, and specifically precludes the analysis of potentially important looping effects.

Looping effects, when understood as a self-categorization process, may be uniquely powerful externalist mechanisms. Indeed the failure of theorists to take external mechanisms seriously, as much as anything reflects theoretical inadequacies in much of what passes as “social” psychology. Ironically, much social psychological research treats social processes ideographically or reduces it to individual differences in demographic characteristics, or individual cognitive biases. Not only does this miss the point that meaning making is a fundamentally inter-subjective process (Turner & Oakes, 1986; Searle, 1995), it provides few insights into how social processes could systematically change the properties of specific HPC disorder categories, rather than serve as generic stressors or buffers. By contrast, the self-categorization account suggested here is potentially powerful, because it provides a theoretical and empirical basis that helps “explain large scale uniformities in behaviour from people’s relations as group members” (Turner & Oakes, 1986 p. 241).

These arguments lead to the upper right quadrant in Figure 5. An unfortunate consequence of this mapping is that it might suggest that if essentialism and HPC internalism are wrong, then HPC externalism must be right and looping effects must explain all of the stability and variation in mental illness. This is not intended. Boyd’s (1989) HPC account, and even Hacking’s account (1999), are agnostic about where causal processes predominate—we should take them where we find them. Some HPC mental illnesses may well turn out to be overwhelmingly sustained by individual differences. For example, there

may be limited lay awareness about a condition, or there may be a very high degree of consensus across lay and expert groups about what a psychiatric category means. But HPC models of mental illness can and should at least test whether psychological groups and group processes are causal mechanisms that influence how symptoms co-occur and cause dysfunction, and in a way that is not reducible to individual differences.

It should go without saying that my account challenges all of the disease model tenets: essentialism, internalism, reductionism, discontinuism and ahistoricism. One interesting point to note is that an HPC account that incorporates looping effects suggests that the taxonomic spectrum proposed by Haslam (2002b)—from practical cut points on a continuum, through fuzzy kinds to discontinuous kinds—may itself be dynamic. After all, if being a Trotskyite is a discontinuous set of beliefs, then it is theoretically possible that looping effects might, under certain circumstances, polarise disorder sufferers in such a way that previously continuous distributions of disorder symptoms may become fuzzy or even discontinuous categories over time.

These dynamics highlight the important point that while HPCs provide a conceptual framework for understanding how symptoms co-occur, this is actually only *part* of the conceptual problem of psychopathology which was to explain not just symptom coherence but also stability and change in mental disorder. The HPC framework could therefore usefully be elaborated to describe not just the fact that the properties of a category co-occur, but that over time there can be variation in: the number of category members, the number and severity of symptoms displayed by average category members, and the distribution of symptoms displayed by category members. In other words, if mental disorders are not timeless essences, but rather categories whose properties can dynamically evolve over time, in interaction with the social and environmental context, then it is important to establish a

more formal framework for describing such changes. This important challenge will be taken up in the next chapter.

I finish this chapter by summarising the core ideas that underpin the alternative analysis of mental illness elaborated so far. The main claims of this revised account are that:

1. Mental disorders are homeostatic property clusters. That is, they are categories, defined by communities of experts, that are constrained by causal structures in the world and which probabilistically predict, explain or control how symptoms co-occur and predict dysfunction in a social environment.
2. Mental disorders are real parts of a complex, dynamic system. They are materially grounded in causal structures that can exist both in and between mind/brains and their environment.
3. Mental disorders may be distributed continuously, fuzzily, or discontinuously in nature, and these distributions may change over time (Zachar, 2000, Haslam, 2014).
4. Mental disorder categories serve different institutional functions. Disorder categories are objects of scientific research, but also serve administrative, judicial, and statistical purposes.
5. Disorder categories may become inter-subjectively meaningful ways of perceiving the self. The set of people who identify with a disorder category may only partially overlap with the people who meet diagnostic criteria (the set of people classed as category members by researchers).
6. To the extent disorder categories are internalised as shared self-categories, HPC disorders may be dynamically shaped by inter-group processes. Beliefs about causal mechanisms that sustain a disorder, the purposes for which a disorder category is used, and the supposed dysfunctionality of “symptoms” can be very different from the perspective of institutional experts and those who self-categorize as “disordered”.

Because of this contestation, the properties of HPC disorder categories can be systematically altered by group processes.

Chapter 2: The Disorder Identity Analysis: the Disorder Identity and HPC Frameworks and Recovery from Mental Disorder

In the previous chapter I argued that mental disorders are best understood as homeostatic property clusters (HPCs). That is, they are categories for explaining how mechanisms cause symptoms to co-occur and predict dysfunction in a social environment. They are also institutional facts because they are constrained by both brute regularities in the world and expert consensus as to how best to operationalise wider social values about dysfunction. The focus of the previous chapter, therefore, was on categorizing mental disorder from the perspective of expert researchers and clinicians. In contrast, in the present chapter I build on Hacking's notion of looping effects to examine mental disorder as a category not just from the perspective of experts, but from the perspective of those *with* mental illness.

Hacking's proposed looping effects suggest that, to the extent people experiencing a psychological disorder collectively perceive themselves as disordered, this may substantively change the properties of the disorder category as observed by experts. But in Hacking's account, the word "change" does a lot of conceptual work. In the present chapter, I therefore attempt to map more systematically the very different kinds of changes that this collective self-awareness can create. To achieve this requires a way of more rigorously distinguishing between clinical expert and first-person categorization schemes, and doing so in a way that allows one to evaluate changes in one scheme from the frame of reference of the other.

In this chapter I therefore build a crucial conceptual foundation for this thesis by describing what I call the HPC Framework (mental disorder categories from the perspective of clinicians) and the Disorder Identity Framework (mental disorder categories from the perspective of those who perceive themselves as disordered). Jointly, these two Frameworks underpin my broader Disorder Identity Analysis. As mentioned briefly towards the end of the

last chapter, the HPC Framework needs to be expanded to include not just homeostatic mechanisms, but also a series of mechanisms that describe the inflow and outflow of category members, as well as changes in the properties of HPC category members.

But to elaborate Hacking's looping kind argument it is also necessary to explain what disorder categories mean from the perspective of those who collectively perceive themselves, in some sense, as disordered. To do this, I develop what I call the Disorder Identity Framework, according to which people who self-categorize in terms of a disorder category can use a shared Disorder Identity to collectively modify both themselves and their social environment. I will claim that it is these, what I call Disorder Identity Adaption Strategies, that drive looping effects and have causal consequences that can be usefully mapped on to the HPC Framework. A key task of the current chapter, therefore, is to develop a conceptual vocabulary for *describing* these processes from the first and third person perspectives. It must be emphasised that the theoretical and empirical evidence for these psychological processes will be elaborated in Chapter 4 and Chapter 5.

But this proposed perspective switching, between the HPC Framework and the Disorder Identity Framework, also raises a perhaps unexpected, but I will argue conceptually important, connection between looping effects and the idea of recovery from mental disorder. I propose that the utility of the Disorder Identity and HPC Frameworks can be understood best by showing how they clarify an often confused and highly contested debate about the nature of recovery from mental disorder. Again, I will argue that this confusion arises in part because of unstated assumptions, informed by the disease model, about what mental disorders are.

I therefore must start by outlining a commonly drawn distinction between cure, recovery-as-an-outcome, and recovery-as-an-orientation, and demonstrate how the usefulness of these concepts differs, depending on whether one presupposes the disease model or the

HPC model. I then explain the conceptual limitations of existing understandings of recovery-as-an-orientation and show why understanding this concept in terms of the Disorder Identity Framework is a more coherent alternative. However, I argue that shared understandings about disorder, from the perspective of disorder identifiers, do not always align with institutional claims about what functioning is and how it is “should” be achieved. Because of this, when viewed from the HPC Framework, these Disorder Identities can sometimes be powerful homeostatic mechanisms, but can in other circumstances help individuals to overcome their symptoms and avoid symptom relapse, and in yet other circumstances can lead to wider social changes that force experts to redraw the boundaries of dysfunction.

Recovery as cure, outcome or orientation?

As briefly mentioned in the previous chapter, the disease model implies that genuine recovery entails cure, where cure just means the disease-causing essence is somehow got rid of or permanently neutralised. As a verb, to cure is transitive—it is something done by an expert *to* somebody or their disease. An important additional criterion for something to count as a cure is that its benefits outweigh harms. The thousands of people who received lobotomies throughout the 20th century certainly had their symptoms permanently neutralised, but few would suggest this “cured” their condition, that the benefits outweighed the devastating consequences (Valenstein, 1986). A cure must therefore be a *targeted* treatment that permanently disables a disease-causing entity.

Compared to cure, the syntax and meaning of recovery is different. The patient is usually the logical subject of a sentence involving the verb “to recover”. Such a sentence implies the patient’s symptoms are at least in remission and, unlike cure, is neutral about how, why or for how long this occurs. However there is a widely drawn distinction between recovery-as-an-outcome and recovery-as-an-orientation (Resnick, Fontana, Lehman, & Rosenheck, 2005).

Recovery-as-an-outcome implies that a patient no longer meets diagnostic criteria defined within the disciplines of psychiatry and clinical psychology. Recovery-as-an-outcome is therefore an institutional fact—much like “mental illness” and “symptom”. Recovery-as-an-*orientation* is importantly different. I will elaborate further what the term might mean, but note for now that it is a process that involves the active involvement of the patient.

Already, these distinctions between cure, recovery-as-an-outcome, and recovery-as-an-orientation, have significant implications in the context of the disease model and the HPC model. If mental illnesses are essentialist natural kinds then, in principle, cures are possible. Palliative treatments for disease essences, such as psychotherapies and psychopharmaceuticals, may lead to symptom remission (recovery-as-an-outcome), but cure, by definition, would involve permanently disabling the disease causing entity. Perhaps surprisingly given the dominance of the disease model in psychiatry, talk of mental illness cure is remarkably rare (Insel & Scolnick, 2006). Indeed Insel and Scolnick, who are both strong proponents of a more biologically-based psychiatry, argue that many have been too quick to assume that psychiatry can at best hope for marginal improvements using the merely palliative treatments that currently exist. Instead, Insel and Scolnick claim “we need to aim for a goal of recovery defined by a complete and permanent remission” (p. 12)—in other words, a cure.

But if mental illness is an HPC kind then aiming for a cure, as defined above, would likely be impossible. If illness is sustained by multiple symptom-generating mechanisms, none of which is necessary or sufficient for disease, then there is no singular locus for a targeted cure, nor a definitive basis for declaring “complete and permanent remission”. Whereas essentialist natural kinds provide nomological certainty about the presence or absence of disease, HPC kinds cannot. Of course, on pragmatic grounds, some may claim

they are “cured”, because they have not experienced symptoms for a long time. But the term “recovered” (as an outcome) more accurately characterises this state and does not presuppose the essentialist baggage and epistemic certainty of cure.

In sum, if the HPC account of mental illness is correct then treatments can support recovery-as-an-outcome by targeting the causal mechanisms that sustain dysfunction. The HPC account does not necessarily have strong implications for recovery-as-an-outcome other than the deflationary conclusion that cures are not possible. But this leaves open the question of what is meant by recovery-as-an-*orientation* and its relationship to recovery-as-an-outcome.

Recovery-as-an-orientation

For clarity and brevity, from now on I use the term “recovery” as shorthand referring to recovery-as-an-orientation and use the term “symptom remission” to refer to recovery-as-an-outcome. In this section, I describe and critique recent attempts to develop a theoretical account of recovery. The critique I develop focuses on a particular account of recovery developed by Davidson and Roe (2007), although the arguments are more broadly applicable to several similar accounts (Anthony, 1993; Jacobson & Greenley, 2001; Resnick et al., 2005). But first, it is important to briefly consider the historical origins of recovery.

The concept of recovery, in the sense used here, began in the late 1980s as a patient-lead movement that sought to challenge dominant biomedical understandings of mental illness. In the wake of widespread de-institutionalisation of mental illness sufferers, former patients redefined themselves as “consumers”—with the implication that consumers have power and choices—and as “survivors / ex-patients”—in part to convey their sense of ongoing trauma at the hands of institutional authorities. The movement particularly challenged the idea that the prognosis for severe mental illness is universally negative, that recovery entails biomedical cure, and that consumers have a limited role in defining and

controlling what a meaningful life entails. A central theme of the movement is that mental health consumers should have greater power to make decisions in their own interests and on their own terms rather than being passive objects of biomedical intervention. The movement has been profoundly influential and has gained widespread acceptance as a guiding principle for modern, western, mental health policy and practice (Hogan, 2003).

But the concept has also been severely criticised: for lacking a coherent definition (Lieberman & Kappellerwicz, 2005; Davidson et al., 2006); for creating unresolvable ambiguity about whether recovery entails cure or symptom remission, or some other, unspecified outcome (Davidson et al., 2006); for setting unrealistic expectations (Masland, 2006); and for lacking a sound evidence base (Remington & Shammi, 2005). Indeed Davidson et al. (2006) argue that the term has been used in different, sometimes contradictory, ways by consumers, advocates, family members and professionals to the point where it seems the only thing these stakeholders do agree on is the lack of consensus about what recovery means.

Notwithstanding, several authors have attempted to articulate a theoretical account of recovery (Davidson & Roe, 2007; Jacobson & Greenley, 2001; Rudnick, 2008). The aim of such theoretical accounts is presumably to provide sufficient conceptual clarity for the notion of recovery to be empirically generative. With this aim in mind, Davidson and Roe offer a fairly typical definition:

Recovery refers primarily to a person diagnosed with a serious mental illness reclaiming his or her right to a safe, dignified, and personally meaningful and gratifying life in the community while continuing to have a mental illness. It emphasises self-determination and such normative life pursuits as education, employment, sexuality, friendship, spirituality, and voluntary membership in faith and other kinds of communities beyond the limits both of the disorder and of the mental

health system, and consistent with the person's own goals, values, and preferences.

(p. 465)

By this account, recovery certainly seems like a universally excellent idea. But consequently, it is hard to imagine anybody, mentally ill or otherwise, who would not seek the right to a “personally meaningful and gratifying life consistent with the person's own goals, values, and preferences”. As Davidson and Roe (2007) concede, a better test of this definition is to ask: how would one know if a person was *not* in recovery? To this question, Davidson and Roe propose three categories of people with mental illness who would not meet their definition:

1. those who have “yet to begin to figure out how to live a meaningful life beyond their disability” (p. 467);
2. those who are overwhelmed by the severity of their mental illness; and
3. those who “choose not to manage their condition” and who “defy all health-care providers' admonitions to take better care of themselves” (p. 468).

These constraints are supposed to provide a more rigorous basis for distinguishing those who are recovering from those who are not.

There are a number of problems with this as an account of recovery. Among them is the lack of any theoretical context for situating recovery, and a resultant lack of clarity about how one might operationalise 1 to 3, let alone generate any empirically testable claims about recovery. But a deeper problem is the inherent ambivalence towards the first person perspective of mental illness. On one hand, recovery researchers and practitioners are at pains to emphasise how recovery is about consumers reclaiming power and defining functioning and meaning on their own terms. On the other hand, there is an attempt to constrain what counts as recovery, based on the values of clinicians and researchers. It is as if reclaiming power and self-determination is a thing to be promoted, but only if it is done

“our” way—it certainly cannot entail activities that “worsen” symptoms. Those who “defy all health-care providers’ admonitions to take better care of themselves” are therefore not considered to be in recovery even if, from the perspective of the consumer, this may be precisely what it means to reclaim power and personal meaning.

Even if these problematic constraints on recovery are accepted, Davidson and Roe’s account can be criticised from the other end—for being too permissive. Davidson and Roe argue their account is attractive because it recognises diverse activities that contribute to what they call the “work of recovery” (p. 464). They argue that seemingly mundane activities, like walking the dog or doing the dishes, may be profoundly meaningful achievements for people with mental illness, and are neglected in narrowly clinical accounts. But their definition provides no grounds for precisely distinguishing whether an apparently mundane activity is, from the perspective of the consumer, actually perceived to be meaningful as a recovery activity. The account can therefore be criticised for not only wrongly *excluding* activities that are perceived to be meaningful (because of criteria 1 to 3), but for wrongly *including* activities that may be irrelevant to a consumer’s understanding of recovery.

A final criticism of this, and similar accounts of recovery, is the near universal assumption that recovery pertains only to particular kinds of serious mental illness—usually conditions such as bipolar disorder and schizophrenia (Resnick et al., 2005; Rudnick, 2008). This may be a descriptively accurate characterisation—in the sense that it captures a particular historical movement led by former patients of psychiatric institutions, many of whom were diagnosed with these conditions. But if putatively theoretical accounts of recovery are more than just redescriptions of a particular socio-historical movement, then they should ideally seek to explain a wider range of phenomena and should accommodate different types of disorder. Some disorders, such as autism, can seriously and profoundly affect the lives of sufferers and their families, but are seldom considered in the context of

recovery. Similarly, it is unclear why common conditions, such as anxiety and mood disorders—which can be either mild or very severe—are largely neglected in the recovery literature.

The analysis above suggests that recovery has been used as a term to describe an historical movement and has been influential in changing the delivery of mental health services. The idea has certainly been well intentioned, and challenges widely held assumptions informed by the disease model, and indeed has served as a powerful rhetorical tool for changing institutional cultures and the expectations of patients, clinicians and families. But attempts to characterise recovery as a theoretical construct have been problematic. Accounts typically have a limited scope, fail to adequately constrain what recovery means, and conflate the very distinction they purport to keep separate: consumer empowerment and symptom outcomes. These problems stem from a broader failure to acknowledge the inherent tension between mental illness symptoms as institutional facts, operationalised from the perspective of clinicians and researchers, and the illness experience from the perspective of patients. Sometimes these perspectives align, but sometimes they do not. This tension is not resolved by arbitrarily stipulating a compromise between clinical and consumer perspectives. Indeed doing so sterilises the very contestability that makes recovery such fertile theoretical ground.

Recovery, adaptation and the self

The foregoing critique implies that a more theoretically robust account of recovery should ideally have greater scope, be empirically generative, and attend more closely to the meaning of disorder and recovery from the first-person perspective. Along these lines, Rudnick (2008) argues that recovery from mental illness can be understood as a compensatory self-organisation process in an environmental context. Since individual functioning is interdependent with both the environment and the person, then adaptations can

involve both external as well as internal compensatory mechanisms. Various theories of disability make similar claims (Cooper, 2012). These accounts are an important advance because they emphasise that there is something distinctive about the first person perspective, about perceiving oneself as being mentally ill, and that adapting to this can involve changing the self as well as the wider social environment.

While these accounts stress the importance of adaptive self-organisation, and the need to “address the person as a whole” (Rudnick, 2008, p. 276), they often neglect the *intersubjective* aspect to this process. Seeing oneself as a depressive or as a mental health consumer is not an idiosyncratic belief that emerges in a social vacuum. Moreover, the intersubjective meaning of mental illness, and the strategies for coping with it, may or may not align with the goals and understanding of institutional authorities or society more generally. Adaptively modifying the social environment is therefore invariably a collective, rather than an individual, achievement.

The Social Identity Approach (Tajfel & Turner, 1979, Turner et al., 1987) provides a powerful alternative analysis of the self for understanding this recovery process. This approach presupposes that psychologically, the self is not always individuated. Instead, the self is understood to be partly constituted by a set of social-category representations of varying levels of inclusiveness. Various “we”-selves can be more or less salient in different contexts and are just as psychologically real as the first person “I” self. Crucially, it is this shift from perceiving the self as an individual, to perceiving the self as a psychological group member (as an Australian, or a conservative, or a depressive), that makes shared meaning and shared behaviour among group members possible.

Defining Disorder Identity

The implication of the Social Identity perspective that I now explore is the idea that recovery as an orientation can be understood as an outcome of psychologically identifying

with fellow members with a disorder. On this view, while it is possible for social categories to be relatively fleeting ways of perceiving the self in a situation, the notion of social-psychological identification with a group implies that a particular category reflects a more enduring and shared self-representation (see McGarty & Grace, 1999, for a more detailed discussion of this point). To the extent that one identifies with a disordered identity, then, one is more likely to *self-categorize* as such in any given situation.

It is also necessary to develop a clearer working definition of what I will call a Disorder Identity. In this thesis, I define a Disorder Identity as referring to a group of people with:

- a shared belief that all category members, either currently or in the past, display(ed) a coherent set of behaviours, feelings, thoughts or sensations
- a shared belief that the set of behaviours, feelings, thoughts or sensations either:
 - o *should* be characterised institutionally (by psychiatrists and psychologists) as clinically dysfunctional, or;
 - o *are* characterised institutionally (by psychiatrists and psychologists) as clinically dysfunctional. Note, this belief could exist even if group members reject that the behaviour is dysfunctional, or disagree with institutional claims about the cause of the dysfunction.
- a shared belief that the group defined by the other beliefs (above) is a subjectively meaningful way of perceiving the self.

Some examples will help make clear which groups would count as Disorder Identities by this definition. A group of people who identify as recovering alcoholics, but who do not meet clinical criteria for any disorder, but nevertheless believe that they and other recovering alcoholics at some stage did, is an example of a Disorder Identity. A group of people who identify as being “pro-ana” and who reject the idea that anorexia nervosa is a disease but

know that typical behaviours of group members would be perceived institutionally as disordered would count as a Disorder Identity. However, a group of people who meet criteria for Alcohol Abuse Disorder, have a shared belief that they drink large amounts of alcohol, but who deny that their behaviour is dysfunctional, and falsely believe that institutional authorities would not perceive their behaviour as particularly pathological, would not count as a Disorder Identity.

A further point to emphasise is that Disorder Identities as defined here are assumed to be a collective phenomenon. This does not preclude the possibility that an individual may perceive his or her experience to be coherent, subjectively meaningful, and considered pathological by institutional authorities, but nevertheless individualistically defined. For example, an individual experiencing acute psychosis may have a radically idiosyncratic understanding of his or her behaviour, but also an awareness that his or her behaviour is defined institutionally as dysfunctional. However, this would not be within the scope of the definition of Disorder Identity as outlined here. The reason for this is that the key rationale for invoking the Disorder Identity construct is to explain large-scale uniformities in psychiatric behaviour. An individualistically defined sense of Disorder Identity, that differs substantially from person to person, would be conceptually uninteresting because it does not provide a conceptual basis for explaining similarities across disorder sufferers. This important point will be elaborated in Chapter 4 where I outline the social-psychological processes underpinning social influence (Turner, 1991) and show how this explains how and why Disorder Identities can form and how these identities can drive shared behaviours.

The definition set out here also implies that the overarching analysis of looping effects is underpinned by two, conceptually distinct perspectives: the first person perspective of the disorder sufferer and the third person “expert perspective”. Importantly, these two perspectives generate four main possibilities as shown in the quadrants of Figure 6.

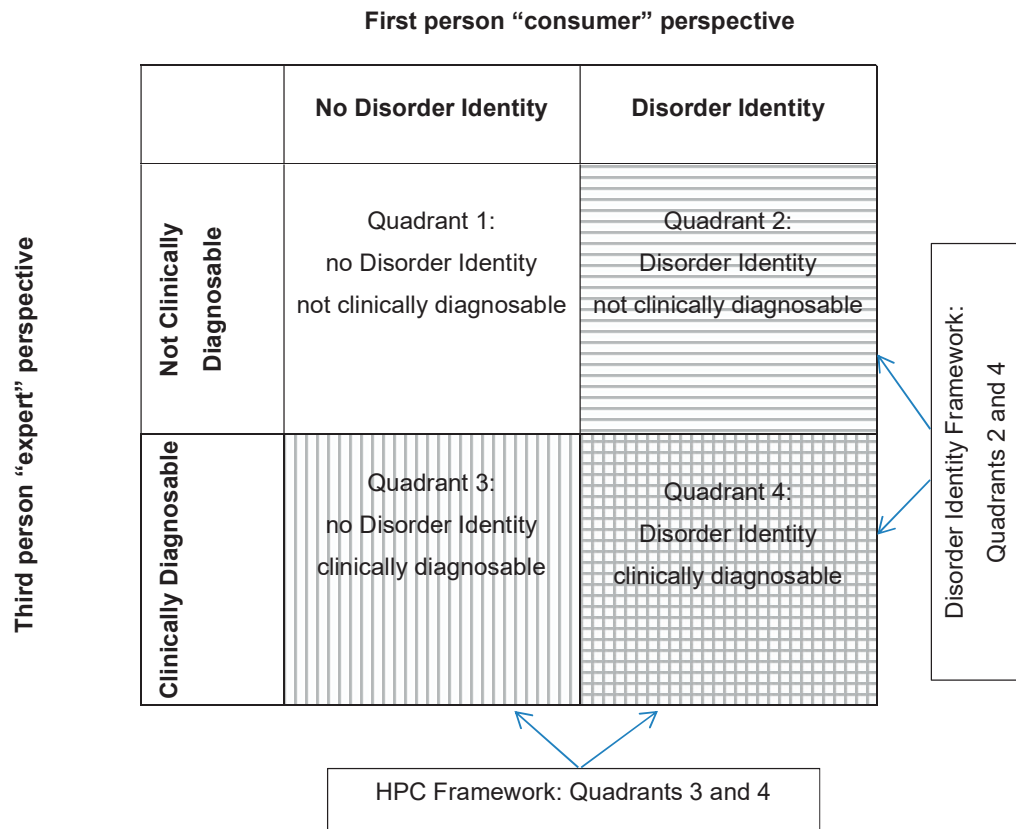


Figure 6: The Disorder Identity Analysis of looping effects: the interaction between the Disorder Identity Framework and the HPC Framework.

The focus of the HPC framework developed in Chapter 1 was on Quadrants 3 and Quadrant 4. Specifically, the HPC framework suggests that one of several causal mechanisms that could generate clinical similarities between individuals with a clinically diagnosable disorder is a shared Disorder Identity (i.e. those in Quadrant 4 above). However, this is neither necessary nor sufficient to define the extension of an HPC category. There may be individuals who meet expert criteria for disorder, in Quadrant 3 above, who are subject to other “brute”, similarity generating mechanisms (unrelated to Disorder Identity processes) that cause them to meet criteria for an HPC defined disorder.

By contrast, the focus of the Disorder Identity Framework is on Quadrants 2 and 4, above. As I will show, people who identify as disordered (whether or not they meet clinical criteria for disorder) will engage in specific behaviours informed by that shared Disorder

Identity. Jointly, these two frameworks underpin the overarching Disorder Identity Analysis of Looping Effects (or, for brevity, the “Disorder Identity Analysis”), which incorporates the dynamic interaction between the first person and third person clinical perspectives implied by Hacking’s “looping kinds” argument.

I further suggest that the Disorder Identity Analysis provides a conceptually more rigorous account of recovery. I argue that recovery-as-an-orientation can be understood as adopting what I call a Disorder Identity Adaption Strategy. This occurs where:

- a psychological group member perceives that a Disorder Identity is self-relevant in a particular situation;
- a psychological group member acts in a way that is informed by a *shared* understanding of what a disorder category means, and the socio-structural constraints of his or her environment.

This definition does not assume that Disorder Identity Adaption Strategies necessarily lead to symptom reduction. In fact, as I will argue later in the chapter, these strategies may actively “worsen” symptoms from a clinical perspective. The definition also rules out cases where a person may well be behaving in a way that alters his or her symptoms, but this is coincidental. The person with Major Depression who goes for a walk and feels better is not an example of a person enacting a Disorder Identity Adaption Strategy—*unless* that person is doing so (at least in part) because they identify as depressive, believe that depression is a disorder and that this is what depressed people can do to relieve their symptoms or live well with depression. This is important. A core sense of the term “recovery-as-an-orientation” I have developed here is the idea that recovery is about acting, rather than being acted on. But on the account proposed above it cannot be just any acting. It is acting as a person who, at that moment, sees him or herself in a particular way: as a person with a disorder.

It is also important to note that often beliefs about the shared meaning of a Disorder Identity, and the constraints of the environment, are tacit and group members' beliefs about the degree of consensus can vary. At one extreme, a person who insists that a Disorder Identity has no coherent or shared meaning, or none that is in any way self-relevant, is a person who is clearly not adopting a Disorder Identity. At the other extreme, especially for Disorder Identity groups with a formal membership structure (e.g. Alcoholics Anonymous), it might be that there is a very clear consensus about what a disorder means and what it is that group members typically do, informed by actual, face-to-face, social interaction. More often, views about the degree of consensus about group beliefs are more tentative. For example, despite not having met any "fellow sufferers" in person, individuals may develop an awareness through various media that there are "people who feel like me". As I will outline towards the end of this chapter, the internet provides an unprecedented platform for people to seek out, discuss, and reach consensus about what a condition means: even without having met fellow sufferers in person.

It should be emphasised that the account above presumes that shared beliefs about the nature of a disorder and the social context will be constrained by "brute" biological and structural realities. Indeed the HPC account of disorder that I developed in Chapter 1 presupposes that psychiatric disorders form coherent syndromes as a consequence of multiple "similarity generating mechanisms": some of which are "brute" and some of which, I argued, are generated by Disorder Identity processes. Group members who propose an adaption strategy that is wildly at odds with these brute constraints will face serious difficulties in maintaining this narrative in the face of competing accounts that can better predict, explain and control group members' lived experience. At the same time, as I shall soon elaborate, the extraordinarily creative ways in which groups can develop quite divergent understandings of the same constraints should not be underestimated.

The term “adaption” used in this context of Disorder Identity Adaption Strategies also needs clarifying. The biological connotations of this term are deliberate but also potentially misleading. Typically within biology, the term adaptive is used to refer to any structure that enables an organism to survive and reproduce. Adaption, in this strictly biological sense, can be explained in a way that does not presuppose particular values or purpose. However, the way that the term has come to be used within western medical and societal discourse *is* value laden. For instance, few would argue that it is maladaptive to choose not to have children or to be homosexual, even if many used to believe this and some continue to. The sense of adaptiveness that is used here is therefore not an empirical claim, but rather a matter of (often contested) social consensus. The valence of a particular condition is therefore a matter of shared agreement within a community whose members identify with a condition, and so the goals for the group partly stem from this shared understanding.

This hints at the biological connotation of adaption that is worth retaining: the notion of ecological niche creation (see Hacking, 1998, for a full discussion of this point). Briefly, this is the now well established idea within ecology that organisms can draw on genetic or acquired information to actively modify their own and others’ environment (Laland & Sterelny, 2006; Odling-Smee, Erwin, Palkovacs, Feldman, & Laland, 2013). Niches, including social niches, can also be passed on to subsequent generations—and are often an intergenerational resource which organisms inherit. Niche construction can therefore have profound consequences on the structure and characteristics of the wider ecosystem—with both positive and negative effects on an individual organism’s biological fitness. The parallels are not explored fully here other than to point out that Disorder Identities are, in a similar way, a *resource* that can be used to maintain or modify the self and the social world and that can be passed on to others.

The Disorder Identity Analysis developed here clearly owes much to Hacking's (1998, 1999) looping kind arguments, discussed in the previous chapter. Like Hacking, the account emphasises the distinction between mental illness from the first and third person perspective. The analysis also emphasises how seeing oneself as part of a mental illness category is in important ways a dynamic and collective process. However, the account is an important elaboration of Hacking's claims, because it provides a more precise articulation of what "collective self-awareness" means and how this could be operationalised empirically, informed, as I will show in detail in Chapter 4, by established social-psychological theory. Moreover, as I discuss presently, this formulation provides a means of keeping track of the diverse kinds of changes that might be predicted to occur as a consequence of adopting a particular Disorder Identity Adaption Strategy.

The consequences of Disorder Identities on HPC categories

So far in this chapter I have developed a definition of Disorder Identity and sketched why defining recovery as adopting a Disorder Identity Adaption Strategy can avoid some of the conceptual problems faced by existing definitions of recovery. The next step is to map out the possible clinical consequences Disorder Identity Adaption Strategies might lead to when viewed from the third-person, clinical perspective (i.e. within the HPC framework). Again, the theoretical and empirical account of the social-psychological mechanisms, explaining *how* these mechanisms work, will be developed in Chapter 4. The point here is to lay the basic conceptual groundwork.

To begin with, I note that, having developed the HPC framework and the Disorder Identity Framework, it is necessary to hold constant the definition of an HPC disorder (remembering, as always, that HPCs are (partly) institutionally defined constructs, and so can evolve over time). This provides a stable frame of reference from which to describe how the properties of members of an HPC disorder category change over time.

To explain this point, consider the fact that to operationalise HPC categories it is necessary to make a pragmatic, institutionally defined cut point about when category membership starts and stops. For example, on an HPC view, the symptoms of Major Depression tend to correlate with each other and this is explained by the presence of various mechanisms that are, to varying degrees, shared by category members. But the HPC account also requires that the discipline makes some judgement about what is going to count as a “case”, for example by stipulating that Major Depression exists if an individual has a Beck Depression Inventory (BDI) score above 13 (as a proxy for how this condition is defined within DSM). However, having made such a stipulation, it is also important to be able to describe different ways that the properties of the set of category members can change. For example, it is useful to be able to describe: mechanisms causing the mean BDI score among Depressives to increase from, say, 15 to 17 over a certain period; or mechanisms that cause the number of HPC category members to increase or decrease over time.

In Table 1, I show that there are at least four broad possibilities that need to be considered and that I define as follows:

1. Homeostatic mechanisms—mechanisms that cause the co-occurrence of particular properties within HPC category members.
2. Homeodynamic mechanisms—mechanisms that facilitate changes in the average properties of HPC category members.
3. Heterodynamic mechanisms—transitional mechanisms that facilitate inflow and outflow of category members.
4. Heterostatic mechanisms—stabilising mechanisms that resist the inflow and outflow of category members.

Interestingly, despite the recent interest in HPC accounts, a systematic theoretical language for describing these processes has not been developed (to my knowledge). I

Table 1: *The HPC framework: homeostatic, heterodynamic, heterostatic, and heterodynamic mechanisms*

1. Homeostatic mechanism 1.1. +Hms = positively homeostatic mechanism, increases the degree of similarity between HPC category members. 1.2. -Hms = negatively homeostatic mechanism, decreases the degree of similarity between HPC category members.	
2. Homeodynamic mechanisms 2.1. +Hmd = positively homeodynamic mechanism, increases the average severity of HPC symptoms. 2.2. -Hmd = negatively homeodynamic mechanism, decreases the average severity of HPC symptoms.	
3. Heterodynamic (transitional) mechanisms 3.1. +Htd = positively heterodynamic mechanism, facilitates outflow of HPC category members. 3.2. -Htd = negatively heterodynamic mechanism, facilitates the inflow of HPC category members.	
4. Heterostatic (stabilising) mechanisms 4.1. +Hts = positively heterostatic mechanism, prevents inflow of HPC category members. 4.2. -Hts = negatively heterostatic mechanism, prevents outflow of HPC category members.	

Table 1 provides a framework ('the HPC framework') for describing, from a third person/expert perspective, the possible outcomes that obtain when Disorder Identifiers adopt Disorder Identity Adaption Strategies. These outcomes are implicit within Hacking's (1999) account of looping kinds. A theoretical and empirical account of how social-psychological mechanisms might generate these outcomes is developed in Chapter 4 of this thesis.

propose that by articulating these processes in this way, it is much easier to keep track of the different kinds of changes that various mechanisms, including Disorder Identity mechanisms, can exert on the characteristics of the set of HPC category members. As I will show, it will be possible to describe more precisely exactly what kinds of changes to institutionally defined disorders can be generated by disorder identifiers.

Hacking (1999), for instance, provides the example of Dissociative Identity Disorder, when the average number of alters increased from 7 to 15 over the course of 10 years. If it could be demonstrated that this was partly caused by individuals psychologically identifying as “split”, then this identity would count as a positive homeodynamic mechanism which tended to increase the number and severity of symptoms experienced by group members. Hacking also notes that the prevalence of DID increased from 200 to 20,000, over this same, 10 year period. In this case, identifying as a disorder sufferer may also have been a negatively heterodynamic mechanism, since it induced individuals who would not have otherwise been (sufficiently) symptomatic to meet criteria for the disorder. Further examples of each of the various mechanisms described above will be discussed towards the end of this chapter, and then the social-psychological processes that could explain these proposed mechanisms will be examined in detail in Chapter 4.

This framework, as it has been presented, assumes that the criteria for diagnosing an HPC-defined disorder are held constant. But of course, Hacking’s looping kind thesis implies that disorder categories may need to be updated to take into account changes in the patterns and consequences of psychopathology, as well as to reflect broader changes in social values about what counts as dysfunction. The dynamic possibilities suggested in Table 1 imply that the properties of HPC disorder categories may, over time, need to be: split apart or lumped together; or thresholds for dysfunction raised or lowered; or even disbanded altogether. Of course many other factors beyond Disorder Identity processes may also

contribute to these nosological decisions. The point is just that, if Hacking is correct, and disorder categories are moving targets (at least partly because of looping effects) then HPC disorder categories will need ongoing revision.

Informed by Hacking (1999), Haslam (2016) has recently articulated a framework for describing a particular kind of conceptual change within an expert discipline, what he calls conceptual “creep”. This manifests in terms of two kinds of expansion of psychological concepts (including clinical disorders). The first kind of conceptual expansion is what he calls “horizontal expansion” – whereby a concept is expanded to incorporate new kinds of phenomena that were not previously considered within the scope of the concept. The second kind of expansion, so called “vertical expansion”, occurs when the threshold for severity required to “count” as a disorder is lowered. These arguments are quite compatible with the account developed here. The current emphasis has been on substantive Disorder-Identity driven changes in the properties of individuals who are classified as disordered, while holding the expert definition of a category constant. However, the full implication of Hacking’s looping kind argument is that the actions of expert classifiers and the people receiving these classifications can mutually and iteratively change each other. Something like Haslam’s account of conceptual “creep” therefore complements the Disorder-Identity driven changes discussed here.

The Disorder Identity framework, coupled with an externalist HPC framework, therefore provides a useful basis for systematically describing the logically possible consequences of adopting Disorder Identity Adaption Strategies. It provides a means for keeping track of the “looping” effects generated between the third person “expert” perspective and the first person consumer perspective. Importantly, this makes explicit the possibilities that Disorder Identities could: increase or decrease the degree of similarity displayed by HPC category members (be positively or negatively homeostatic); could

amplify or attenuate the symptom severity experienced by the average category member (be positively or negatively homeo-dynamic); could facilitate transitions in and out of an HPC category (be positively or negatively heterodynamic), or; could resist inflow or outflow of HPC category members (be positively or negatively heterostatic). The framework will be used in later chapters to underpin a set of empirical questions about how Disorder Identities are constructed, how they become psychologically internalised, and how the meaning of a Disorder Identity (from the perspective of the group of Disorder Identity members) is interactively defined by the social context. In turn, empirical predictions can be made about how the varying meanings of a disorder (from the perspective of Disorder Identifiers), and varying adaption strategies, may have either homeostatic, homeodynamic, heterostatic, or heterodynamic consequences for a given HPC disorder category.

Disorder Identities in action

In this section, I illustrate the varying ways that individuals may come to adopt a Disorder Identity and how the group's shared beliefs might alter their adaptive strategy. Again, my purpose here is not to elaborate the psychological processes that enable these identity dynamics (this will be developed in Chapter 4), but rather to demonstrate the utility of the conceptual framework in systematically describing the possible consequences of Disorder Identities. Perhaps the most straightforward example of adopting a Disorder Identity relates to the social dynamics of seeing a doctor. People typically go to the doctor with an expectation that that professional provides them information about what is wrong with them. That is, information about who they are. Historically, this may be as minimal as informing people that they are sick. Parson's (1951) sick role theory elaborates this process. He argues that in western countries at least, doctors are socially sanctioned to determine whether people are sick, and if they are, then there is an acceptance, and indeed an expectation, that patients will adopt the "sick role". This entails the responsibility to comply

with the doctor's treatment regime and, in exchange, patients may relinquish their normal social duties. Being "sick" is therefore a powerful, albeit not very specific, Disorder Identity for adapting to an environment: it outsources detailed knowledge about what a condition is to a group of socially sanctioned experts and allows sick people to explain and justify, to themselves and others, who they are and how they behave (rest, take medication, return to their normal social role when symptoms resolve).

The sense of intersubjective awareness about what it means to be sick can therefore be a tacit understanding. It does not necessarily require knowing or interacting with others who are sick, just a socially shared understanding that this is what sick people typically do. It is a role that is generally well understood, and about which there is often a high level of consensus. Depending on the nature of the condition, the sick role may be a Disorder Identity with powerful, positively heterodynamic consequences. This is because, if it is a condition for which evidence-based treatments exist, then identifying with the sick role—understood, in part, as a role that requires complying with a Doctor's advice—is a way of categorizing the self that tends to facilitate *the transition out of HPC disorder*.

As the ethos of patient-centred care has gained traction, clinicians are increasingly likely to provide more detail than merely a generic "sick role"—for example, by making and sharing a formal diagnosis ("you are depressed") as well as a clinical formulation about what this category means. Again, this can be seen as an attempt to persuade the patient that this way of self-categorizing validly applies and usefully explains an aspect of who the person is. Without delving yet into the substantial literature on social influence (Turner, 1991), the significance of this process is just that clinicians have a socially sanctioned role of communicating and influencing individuals that they *should* see themselves in a particular way. This is underpinned by a network of beliefs that are, to varying degrees, shared

between doctors and patients—that medical science is valid and that doctors are legitimate practitioners of it.

Indeed various therapies are specifically designed to foster this acceptance. Those who wrongly perceive benign bodily signs as symptoms are even proffered a special category—health anxiety—to better explain who they really are. Those who deny that a category applies to them may be offered various motivational therapies. For example, the trans-theoretical model (DiClemente & Prochaska, 1998) and other stage models, such as the 12 step program (Ferri, Amato, & Davoli, 1996), in different ways try to encourage consumers to accept who they are: “my behaviour is dysfunctional”, “I *am* an alcoholic”.

The consequences of perceiving a category as self-relevant or not are varied. Cold-sufferer who do not see themselves as such are an annoyance to colleagues and friends, but their symptoms resolve regardless. The diabetic or asthma sufferer who rejects these categories as self-relevant may risk their life. Qualitative research by Adams, Pill and Jones (1997) found that those who psychologically identified with the social identity, “asthmatic”, tended to pro-actively monitor their condition and comply with their treatment. By contrast, those who knew they had been diagnosed as asthmatic but dismissed the diagnosis and viewed their condition as mere breathlessness, were less likely to regularly take their medication and tended to experience more difficulties managing their symptoms. Adams et al.’s research therefore suggests that, in the context of a chronic condition requiring long term management, a Disorder Identity can powerfully attenuate symptoms (i.e., be a negative homeodynamic mechanism)—even if it does not lead to cure.

Adams et al. (1997) also note that merely perceiving an illness-related category as self-relevant, alone, is manifestly insufficient in determining an adaptive strategy. What matters is an individual’s beliefs about what an illness category means in a social context. At a minimum, accepting the sick role means doing what the doctor says and resting until the

doctor declares one fit to return to one's normal role. More contemporary clinical norms encourage practitioners to provide detail that is sensitive to a patient's circumstances. For example, in the context of mental health, almost all validated therapies include a component of "psycho-education" that explains various manifest and dispositional properties pertaining to a disorder category such as the aetiology, course, chronicity, and prognosis. It is this theoretical information about what a category means that shapes expectations and strategies for adapting to an environment: be that seeking to change symptoms through various psychological or pharmaceutical therapies (as in an episode of depression), or finding ways to live well in spite of ongoing symptoms (as in chronic pain management).

Some might argue that psycho-education, and therapy more generally, should be based not on categories, but on the unique and individual characteristics of the consumer. More generally, many advocates of the recovery approach argue that recovery must be understood as a deeply personal, individualistic process (Anthony, 1993; Rudnick, 2008). Taken literally, this claim is implausible. It implies a radical particularism that, far from elucidating the meaning of illness from the patient's perspective, is antithetical to it. While categories, be they folk or scientific, do not provide the law-like certainty of essentialist natural kinds, they do probabilistically predict, explain or control how category members respond in context (McGarty, 1999). To embrace radical individualism is to abandon the possibility that the experience of similar others has any bearing on how a person might respond to his or her situation (see also Porter & Zachar, 2012, who make a similar point). What advocates of individualism are more likely objecting to is not category use *per se*, but rather how and when categories are used. This important issue will be elaborated more fully in Chapters 3 and 4, but for now, the point is just to claim that the process of adopting a Disorder Identity Adaption Strategy depends on an awareness that there are relevantly similar others whose experience helps predict, explain or control the self.

The idea that disorder meaning is an inherently collective and often contested process is vividly illustrated by research examining online disorder-based groups. This technology enables unprecedented opportunities to collectively mobilise, especially for people with low prevalence and concealable disorders. It is therefore a powerful technology for connecting what Solomon (2012) calls “horizontal identities”—identities that are not passed down within families or even particular cultures (vertical identities), but rather diffuse and distributed across individuals who share some common experience. A national survey of over 3000 Americans, for example, found that in the last 12 months about one third of Americans went online to access health information, about one quarter read about somebody else’s personal experience of a health or medical problem, and approximately 16 percent tried to find others who might share the same condition as them (Fox & Duggan, 2013). Just one of the many hundreds of thousands of global online mental health support forums boasts 400,000 members and over two hundred online support groups for people with conditions ranging from Dissociative Identity Disorder (with over 73,000 posts) to depression (with over 200,000 posts). The significance of finding *similar* others is demonstrated by the online, patient-led website called “Patients Like Me” (Frost & Massagli, 2009). The website has over 300,000 participants, and encourages members to “compare treatments, symptoms and experiences with people like you” and to “share your experience, give and get support to improve your life and the lives of others”. These observations suggest that individuals do not just passively perceive Disorder Identities in the world, they may actively seek them out.

Countless other real world “self-help” movements, each with varying degrees of organisation and endorsement from institutional authorities, claim to help people to either change their symptoms or adapt their lives to live well in spite of symptoms. Again, many of these groups promote very detailed theories about what a disorder category means and the implications for consumers. Traditional AA programs, for example, contend that not only

does recovery involve accepting that one is an alcoholic but that alcoholism is an *essential* disposition (a disease) for which there is no cure. On this view, members are encouraged to accept that one can at best hope to be a dry drunk, by committing to the group for life and mutually supporting fellow group members to avoid relapse. This implies that a person may no longer meet criteria for substance abuse disorder (as defined institutionally), but may nevertheless psychologically identify as a recovering alcoholic. In theory, it is because the AA identity has an essentialist understanding of alcoholism, and so motivates ongoing vigilance and self-monitoring, that the AA identity helps prevent relapse. To the extent this is true, identifying as an AA sufferer would therefore be classed as a positively heterostatic, HPC mechanism, because it inhibits the inflow of members into the HPC category of Substance Use Disorder.

The examples provided so far are of adaptive strategies that more or less align with social and institutional consensus about what counts as dysfunction and how it can best be achieved. Some (Bell, 2007) are quick to point out that various patient-led self-help groups can have unintended consequences and some groups are more evidence-based than others. But again, from the perspective of consumers, the identities these groups foster are plausible strategies for adapting to an environment and are strategies that largely accept dominant social and institutional views that symptoms are pathological and should be either managed, endured or overcome.

But it should not be surprising that communities of consumers may adapt, resist or react to the mental illness category meanings and adaptive strategies proffered by institutional authorities. Most famously, gay and lesbian people acted collectively to resist dominant psychiatric opinion in the mid-twentieth century and ultimately achieved great success in making homosexuality not just non-criminal and non-pathological, but a legitimate and proud way to be (to varying degrees in western cultures at least). As a form of adaptation, this was

distinctly different from other adaption strategies. It did not entail changing or adapting the self directly—for example, by attempting to cure homosexuality, or by attempting to palliate ongoing homo-erotic “symptoms”. It involved radically rejecting the idea that homosexuality was an illness and that homosexual acts and thoughts are symptoms. Up until 1973, then, an identity as a gay rights activist would have likely counted as a causal mechanism that amplified symptoms for the HPC kind “homosexuality”—a sexual disorder listed in DSM-II. But of course, this identity activism radically changed both medical and wider public understanding of the disorder. As a Disorder Identity Adaption Strategy it was an important step towards modifying the social environment so that homosexuality was increasingly accepted, societally and eventually institutionally, as part of normal human variation, and therefore no longer a disorder category.

There is no reason to believe this was a one-off. Recently, DSM-V revised the language of DSM-IV-TR, which listed transgenderism as “gender identity disorder” within “Sexual Dysfunctions and Paraphilias”, and listed “Gender Identity Dysphoria” in its own section to “remove the connotation (in DSM-IV-TR) that an individual is ‘disordered’” (see APA, 2013, p. 1). A number of currently active consumer-support groups aim to achieve similar institutional reforms. These groups challenge, to varying degrees, wider societal and institutional assumptions about what constitutes dysfunction. The “hearing voices network”, for example, provides a forum for people who experience auditory hallucinations to share their experience in a supportive environment and seeks to normalise and validate their experience (Corstens et al., 2014). Again, the internet has spawned myriad online groups that more actively reject psychiatric and social orthodoxy and even celebrate the positive virtues of a condition. The neurodiversity movement advocates that Dyspraxia, Dyslexia, Attention Deficit Hyperactivity Disorder, Dyscalculia, Autistic Spectrum, and Tourette Syndrome should be considered part of normal human variation, that “neurotypicals” just do not

understand (McWade, Milton, & Beresford, 2015). The Autism Rights Movement, for example, argues that autism is not a disorder, that attempting to cure it is offensive and that “treatments” for it are often unethical (Jaarsma & Welin, 2011).

Again, these communities foster very particular theoretical understandings about the meaning and causes of a condition. The Fat Acceptance Movement emphasises the genetic causes of obesity—with the implication that those who experience binge-eating disorder cannot be held responsible for their condition (Kwan, 2009). Conversely, online pro-anorexia forums actively encourage members to sustain an extremely low Body Mass Index (BMI). Here, the normative understanding about what the category means is importantly different: the claim is that being thin is a positive *choice*, a sign of beauty and self-discipline, and with the implication that gaining weight is a *moral* failure (Riley, Rodham, & Gavin, 2009). That clinicians and society at large deeply reject this understanding of anorexia is not at issue. There is ample empirical evidence that maintaining a very low BMI is potentially fatal. What is at issue is the ensuing value judgement about whether such a way of being is legitimate or not. No empirical observation will answer this question beyond pointing to the overwhelming societal and clinical consensus that anorexia is profoundly dysfunctional and is therefore a disease to be treated.

Recovery: summary and conclusions

The foregoing examples show how narrowly focused much of the clinical debate about recovery has been. By expanding the scope of enquiry, the Disorder Identity Analysis developed in this chapter shows the diverse and creative ways that people “recover” from disorder categories and use them to modify themselves and their environment. Importantly, the account avoids two major conceptual problems that have bedevilled attempts to empirically characterise recovery to date. First, it avoids the problem of classifying any volitional behaviour that does not actively worsen symptoms as recovery. On this view,

individuals who engage in various activities of daily living, and do not actively defy the recommendations of healthcare workers, are considered to be recovering (Davidson & Roe, 2007). It is a judgement that could be made from a third person perspective. The alternative developed here ties the recovery process to intersubjective beliefs about the existence of a disorder-related category, and the belief that that category is *self*-relevant to the consumer. On this view, recovery behaviours are just those where the person is acting in a way that is informed by shared beliefs about what a category means and the constraints of the environment.

The second advantage of the Disorder Identity Analysis is that it more clearly distinguishes the consumer and clinical perspectives. By contrast, standard accounts of recovery typically attempt to strike a middle ground between two, competing normative intuitions: empowering consumers to define what functioning is and how to achieve it on their own terms; and insisting that there are some things that are somehow objectively dysfunctional or maladaptive. The Disorder Identity Analysis attempts to be more transparent about where the values that inform judgements of functionality come from, and sets up the empirical problem of explaining how and when Disorder Identities become prepotent, how various agents (including clinicians, consumers and society at large) influence each other about what these categories mean and the clinical consequence of internalising these understandings.

The analysis here also sheds light on the relationship between disorder HPCs (disorder categories from the perspective of clinicians), Disorder Identities (disorder related categories from the perspective of consumers), and looping effects (their interaction). The account suggests that HPCs and Disorder Identities are not necessarily antagonistic, as might be implied by Hacking's (1999) account of looping effects. Under some circumstances, Disorder Identity Adaption Strategies may powerfully attenuate causal mechanisms that

sustain an HPC disorder. For example, there is some evidence that being in AA significantly helps to maintain sobriety (Ferri et al., 1996)—even if one of the core tenets of AA (that alcoholism is an essentialist natural kind) is, in all likelihood, empirically false. Other adaption strategies are likely, from a clinical perspective, to be HPC causal mechanisms—the norms associated with pro-anorexia identities are likely to be a causal factor that amplifies anorexic symptoms (positively homeodynamic), facilitates transition into the category (negatively heterodynamic), and serves as a barrier to leaving the category (negatively homeostatic). Importantly, these consequences can help guide clinicians in their attempts to more sensitively understand what illness behaviours mean from the perspective of consumers, while at the same time understanding the clinical consequences of these behaviours within an HPC framework.

But the HPC account also demands an acceptance of the provisional and perspectival character of HPC disorder categories. The vicissitudes of the category “Sexual Orientation Disturbance” is a poignant reminder that any attempt to classify some observed regularity in the world as a disease is inherently value laden. Because there is such strong consensus that many bio-medical disorders are dysfunctional, it can be tempting to presume that pathology can be straightforwardly “discovered” empirically, in the same way that neutrinos or black holes are discovered. Rejecting this assumption does not inevitably lead to anything-goes relativism or necessarily suggest an antagonism towards empiricism. Rather, it demands an acknowledgement that categorization is a flexible process that depends on both brute processes in the world, group-based judgements about purpose and value, and their ongoing interaction. On this account, the environment is not a constant background factor to which organisms passively, and to varying degrees, adapt. Rather, the environment, especially the social environment, is iteratively modified, bequeathed to subsequent generations, and further modified. These dynamic processes generate novel environments and phenotypes and

frequently frustrate attempts to articulate timeless truths about regularities in psychiatric behaviour. But this sense of frustration says more about our epistemic expectations than about the problems with abandoning essentialist accounts of mental illness and recovery.

One potential criticism with the foregoing analysis is that it is disingenuous to equate Disorder Identity Adaption Strategies with “recovery”. It seems at best confusing to describe somebody, for instance, who identifies with a pro-anorexic group as “recovering”—without making more explicit that the sense of the term being invoked is akin to “reclaim”, rather than the more common, lay, understanding of symptom remission. There are arguably more insidious consequences. By eliding these two very different meanings, such an account potentially legitimises dangerous and extreme behaviour.

These concerns are understandable but perhaps just illustrate how fraught it is to proclaim what recovery “really is”. Recovery will no doubt continue to play an important and contested rhetorical role in claims by clinicians and consumer groups. But as many have noted, recovery has become an idea so widely and indiscriminately used that researchers interested in understanding the social dynamics of disorder might be better served by abandoning it altogether. The foregoing analysis hopefully explains why it *is* important to understand the consequences of identifying with a disorder category. What these processes are called is not crucial to the argument developed here. Rather, the point is to provide a theoretical account that allows one to switch between the frame of reference of clinicians and the frame of reference of those with disorder. The externalist HPC model developed in the previous chapter provides a relatively stable perspective from which to evaluate the consequences of adopting a Disorder Identity. Conversely, from the frame of reference of those who identify with a disorder, HPC categories may be the very things that disorder identifiers are seeking to change. Attempting to analyse these dynamics by stipulating in advance what recovery “really is” obscures these fundamentally inter-group processes. By

contrast, analysing this as a contested, social-psychological process of categorization—from the perspective of clinicians (as HPCs) and from the perspective of Disorder Identifiers (as self-categories)—brings these dynamics to life.

A further possible criticism of the account developed above is the neglect of instances where people overcome their symptoms without adopting a Disorder Identity. For instance, a growing literature suggests that participation in various membership groups (with no links to disorder or overt therapeutic aim) can positively benefit mental health (Cruwys et al., 2013). In response, it should be noted that many disorder symptoms resolve without sufferers necessarily being aware or believing they were disordered. It is also an open empirical question as to whether, assuming symptom remission is a desirable goal, particular Disorder Identities, or Disorder Identity Adaption Strategies, are things to be encouraged. For reasons that should now be clear, there will be no single answer to this. Again, the issue that is theoretically important is to understand the perspective of groups of people who are behaving, thinking and feeling in particular ways because they believe some disorder category applies to them.

Finally, the examples developed here are just that. Their purpose has been firstly to sketch how consumers might come to collectively understand who they are and what their putatively disordered behaviours mean in an environmental context. Secondly, the account allows for a useful conceptual synthesis and integration of HPC accounts of mental disorder, looping effects, and the recovery literature. The next step is to elaborate and organise the social-psychological processes that can help explain how Disorder Identities become internalised, how the meaning of Disorder Identities depends on the socio-structural context, and the consequences of this for how consumers think and act.

Chapter 3: Social-Psychological Processes in Mental Disorder

In the previous chapters I explored various meta-theoretical positions that underpin empirical approaches to disorder and recovery. I also developed and defended a particular theoretical understanding that helps to make sense of the dynamic interaction between categorizers and the categorized. But as yet, I have not reviewed in detail how this perspective aligns with social-psychological evidence about the causes and consequences of internalising labels and how varying lay understandings are likely to shape mental disorder symptoms.

In the first section of this third chapter I explore how, according to dominant social-psychological and sociological models of labelling, disorder labels are ultimately disempowering and that on balance consumers would probably be better served by disbanding labels altogether. These perspectives contend that disorder labels are invariably internalised as part of the self-concept and that this in turn causes consumers to passively and automatically conform to a static, negative stereotype of “being mentally ill” and inevitably leads to discrimination and prejudice from the non-mentally disordered mainstream. Far from helping consumers to adapt to their environment, these arguments suggest that labels invariably worsen symptoms because of self-stigma and stereotype threat. This seems to present a serious challenge to the propositions I developed in the previous chapter. However, I argue that these claims are conceptually confused and presuppose a conceptually problematic understanding about the nature of stereotypes, as even some of the proponents of these positions are starting to concede (e.g., Corrigan and Watson, 2002). Because of this, I argue, these perspectives do not pose a threat to the analysis developed in the earlier chapters.

In the second section of this chapter, I examine the social-psychological evidence relating to lay theorising about disorder. This is important because in the previous chapter I argued that the shared meaning of disorder was crucial to understanding how the

consequences of Disorder Identities play out. Here, I argue that, consistent with this, empirical research examining lay understandings of mental disorder provide important insights into how people with mental disorder understand their own behaviour and are treated by others. In turn, this has significant implications for motivational processes in recovery strategies. However, I conclude that much of this theorising views the underlying social-psychological processes as operating at the individual and interpersonal level. Consequently, research investigating lay understandings of disorder neglects, and therefore struggles to explain, the dynamic, inter-group processes that I have begun to explore in the two previous chapters.

In sum, in this chapter I have two main aims. Firstly, I defend the Disorder Identity Analysis I developed in the previous chapter from a set of potentially serious criticisms, informed by social psychological evidence, that suggests internalising Disorder Identities cannot, in any sense, be considered adaptive. In the second part of the chapter, I review social psychological evidence showing the profound importance of lay theoretical understandings of disorder in shaping people's motivational responses to disorder. Although much of this evidence is individualistic, it does provide important insights that will greatly inform the Social Identity based account of Disorder Identity that I develop in the next chapter.

Internalising illness labels: labelling theories, self-stigma and stereotype threat

Labelling Theory

There is a long and rich sociological literature about the consequences of internalising illness labels (Goffman, 1963; Parsons, 1951; Scheff, 1970, 1974). To varying degrees, these theories adopt a symbolic interactionist meta-theory, first articulated by Mead (1932), according to which human social interaction determines how people interpret the world and their sense of self. Symbolic interactionism is a meta-theoretical framework that emphasises

how the self produces and is produced by the social world, and contends that human behaviour is determined by the symbolic meanings and roles that arise from this interaction. The most developed early application of this framework in the context of mental illness is Scheff's (1970, 1974) labelling theory of mental illness. According to labelling theory, western societies developed the idea of mental illness as a way of explaining and controlling behaviours that deviate from cultural norms. Scheff further argues that culturally shared stereotypes about the nature of mental disorder become ingrained from a young age and are perpetually reinforced in the media and in social interactions more generally. It is these negative cultural stereotypes about mental disorder that underpin the negative consequences of disorder labels.

Scheff argues that when individuals deviate from social norms, for example by displaying bizarre, irrational or violent behaviour, they are often brought to the attention of institutional powers, including medical authorities and the criminal justice system, and they are subsequently labelled as mentally ill. He then argues that stereotypes about the mentally ill become the "guiding imagery for action" (1963, p. 447) for both the labeller and the labelled. The labelled individual is often distressed and susceptible to suggestion, and it is invariably easier to conform to the sick role (Parsons, 1951) than try to contest it. Meanwhile, this role conformity is further reinforced by the expectations of expert clinical authorities and others—including family members, employers, and the judiciary. In this way, mental illness emerges from the interaction between individuals and their wider social context, rather than being a consequence of some pre-existing attribute of the person. Labelling theory is therefore a provocative etiological hypothesis claiming that deeply held cultural stereotypes about mental illness, reinforced by powerful institutional systems, cause mental illness.

There are many criticisms of Scheff's arguments on theoretical (Gove, 1975, 1980) and empirical grounds (Whitt, Meile, & Larson, 1979). Most of these critiques amount to a meta-theoretical rejection of what is assumed to be Scheff's strong externalist claim that mental illness can be explained as a sociological process and that labels cause mental illness, rather than the other way around. To be fair, Scheff (1963, 1974) emphasises his hope for a synthesis of sociological and medical models of mental illness. But this points to a tension in his work. On one hand he implies that disorder labels mostly do harm by creating mental illness symptoms, but on the other hand he suggests that a diagnosis may legitimise the mentally ill role as a means to access respite. Moreover, in the moment when individuals are brought to medical attention, being mentally ill is often preferable to the potential alternative labels—such as being a moral or criminal deviant. Scheff's account is therefore actually compatible with the suggestion that, at least in the short term, identifying as mentally ill might be an adaptive way to negotiate social reactions to deviancy.

But the scope of Scheff's labelling theory is limited to the consequences of involuntary commitment (see Thoits, 1985). The endpoint of this commitment, according to Scheff, is that labels calcify into a deeply ingrained illness identity and stable mental illness symptoms. According to Scheff, it is this passive conformity to the mentally ill role that underlies the insidious consequences of labels in the long term. But this characterisation of mental disorder is not born out empirically. It provides few insights into how people overcome mental disorder symptoms, as the majority of sufferers of mental disorder do (Harrison et al., 2001). It is also unclear why individuals often voluntarily seek out diagnosis and treatment, particularly for putative instances of milder and more transient mental disorder (Thoits, 1985).

These latter objections reflect a broader problem that Scheff's theory implies the mentally disordered automatically conform to a monolithic stereotype of "being mentally ill".

This raises two issues. First, such a generic stereotype, as an account of mental illness, does not adequately address how the very specific constellations of signs and symptoms seen in particular disorders emerge. Second, the mechanisms that lead to conformity are vaguely specified, an objection that has been raised in relation to symbolic interactionist approaches more generally (Hogg, Terry, & White, 1995; Turner, Oakes, Haslam, & McGarty, 1994), such as role theory and identity theory (Stryker & Statham, 1985). In particular, the idea that individuals unconsciously and passively conform to relatively stable role stereotypes is challenged by evidence showing how the normative content of the role, and associated behaviours, can change very rapidly as a function of contextual change (Haslam, Oakes, Reynolds, & Turner, 1999; Haslam & Turner, 1992).

In sum, Scheff's labelling theory was a theoretically important early attempt to understand how the self in general, and mental illness in particular, emerges from the social context. However, without elaborating how deviant labels are internalised and how the meaning of these labels can dynamically, and often rapidly, respond to changes in the social environment, the approach risks degenerating into a socially deterministic account. In part for this reason, labelling theory has received limited attention in more contemporary analyses of disorder labels.

Modified Labelling Theory

To avoid some of these objections, Link et al. (1989) proposes a more modest, so-called modified labelling theory, according to which mental disorder labels negatively affect self-esteem and exacerbate pre-existing mental disorder symptoms and vulnerabilities. Link et al. thus rejects the strongly externalist position that labels directly cause mental illness. Instead, Link et al. argue that labelling exacerbates prior mental illness because of the consequences of mental illness stigma. Link and Phelan (2006) build on Goffman's (1963) classic sociological definition of stigma as "an attribute that is deeply discrediting" (p. 3), by

elaborating how stigma is a process involving five inter-related components. The first two components of stigma involve firstly labelling of differences and secondly “the process of stereotyping in which the labelled person is linked to undesirable characteristics” (Link & Phelan, 2006, p. 528). In describing these first two components, Link and Phelan explicitly draw on core ideas within social cognition (Fiske & Taylor, 1991) by arguing that categorization and stereotyping are largely automatic processes that trade accuracy in favour of cognitive efficiency. As a result, victims of stigma are often unfairly judged to display negatively stereotyped characteristics, rather than more accurately judged according to their attributes as unique individuals. The third component of Link and Phelan’s account of stigma relates to categorization—how labellers draw a distinction between “them” (the labelled, negatively stereotyped group) and “us”, thus implying a social distance between the labelled and the labellers. According to the fourth component, because the labelled are set apart and linked with negative characteristics, stigmatisers will often develop a narrative to justify discriminatory attitudes and treatment towards the stigmatised group. The final component of stigma is that the stigmatised group must be disenfranchised. So, even though the mentally ill may engage in components one to four, for example by unfairly characterising their doctors as “pill pushers”, doctors are not a stigmatised group because they remain in a position of social power.

Link’s definition of stigma thus encompasses a complex set of phenomena, but one of the critical elements that causes the negative consequences for the stigmatised is discrimination—that is, unjustified or negative treatment on the basis of a sufferers’ membership of a social group. According to Link, this can take a number of forms. First, negative societal stereotypes towards the mentally ill can lead to both direct and structural discrimination and this, in turn, has demonstrated negative effects on: social interactions and social networks (Harris, Milich, Corbitt, Hoover, & Brady, 1992; Lennon, Link, Marbach, &

Dohrenwend, 1989), housing and employment opportunities (Farina, Gliha, Bourdreau, Ale, & Sherman, 1971; Schulze & Angermeyer, 2003; Wahl, 1999), depression (Link, 1987), and quality of life and wellbeing (Rosenfield, 1997). These negative effects of discriminatory behaviour towards the mentally ill can also persist after successful treatment. Link, Struening, Rahav, Phelan, and Nuttbrock (1997), for example, showed that even after controlling for prior symptomatology, patients who had been treated for and largely recovered from mental illness and substance abuse continued to experience discrimination and rejection (because of their psychiatric history) and this explained up to 10 per cent of the variation in their residual depressive symptomatology.

This evidence demonstrating the negative consequences of received stigma is important and cannot be dismissed lightly. But what is perhaps unclear is how modified labelling theory *explains* the cause of this stigma. The social-cognitive account implies that this negative treatment is a nearly inevitable consequence of human cognitive processes. Indeed negative stereotyping and discrimination would seem to be historically invariant, so long as people continue to use social categories and people with mental disorder have low power. But the history of inter-group relations is replete with examples where power relations between groups radically change over time. Few would argue that the gender wars are over, but it is nevertheless clear that the received stigma of being a woman or gay in contemporary westernised societies, compared to just fifty years ago, is profoundly different. On the modified-labelling account the only way this could be possible is because of changing power dynamics. But modified-labelling theory provides few if any insights about how power relations come to be as they are and how they change. Consequently, by under-theorising the nature of power dynamics, and emphasising the near inevitability of human cognitive processes causing stigma, the theory actually risks perpetuating the very stigma it (implicitly) seeks to expose and challenge.

But modified labelling theory proponents contend that disorder labels are not just damaging because they cause *received* stigma. Link (1989) argues that perhaps the most insidious consequence of labels is that they often lead to so-called *self*-stigma—whereby the stigmatised internalise mental illness as part of their self-concept and come to believe that negative stereotypes about the mentally ill apply to themselves. According to Link, this occurs because those experiencing mental illness are likely to be aware of negative stereotypes about the mentally ill and so, when given a disease label, they are likely to believe these attributes apply to themselves and so experience low self-esteem and low self-efficacy.

On the face of it, this proposition is a serious challenge to the claim that internalising a disorder label is, from the perspective of a self-labeller, in any sense adaptive. It is also a proposition that is seemingly supported by an extensive body of empirical evidence. A meta-analysis by Livingstone and Boyd (2010), summarising evidence from over 45 studies, finds that self-stigma is robustly linked with poor self-esteem ($r = -.58, p < .001$), lower empowerment ($r = -.52, p < .001$), and reduced self-efficacy ($r = -.55, p < .001$). But there are a number of serious problems with the way self-stigma is operationalised in these studies and indeed the concept of self-stigma more generally. One common definition of self-stigma used in these studies is that it is “the loss of self-esteem and self-efficacy that occurs when people internalise the public stigma” (Corrigan, Kerr, & Knudsen, 2005, p. 179). But it is little wonder that self-stigma by this definition is highly correlated with low self-esteem and low self-efficacy. Even if one accepts the causal inference from the putative correlation between self-stigma and low self-esteem, this “empirical discovery” seems to amount to the vacuous proposition that the loss of self-esteem and self-efficacy *causes* low self-esteem.

This conceptual confusion becomes more apparent when purported measures of the construct are examined in detail. The most common measure of self-stigma within the meta-

analysis by Livingstone and Boyd, one which was used in 21 of the 45+ studies, was the Devaluation – Discrimination Scale (DDS) (Link, Mirotznik, & Cullen, 1991). This is a 12 item measure which assesses the extent to which respondents believe that people with mental illness will be devalued or discriminated against. But, as a number of authors have noted, this is actually a measure of anticipating *received* stigma, certainly not self-labelling, nor beliefs about what the label means (Corrigan et al., 2005).

The second most common instrument in the meta-analysis by Livingstone and Boyd (2010), used in 10 of the studies, was the Internalised Stigma of Mental Illness Scale (ISMIS) (Ritsher, Otilingam, & Grajales, 2003). This is a 29 item questionnaire with five subscales measuring alienation, stereotype endorsement, social withdrawal, and stigma resistance. Again, none of the items have a clean measure of whether people perceive that a mental illness label is self-defining. Moreover, the authors report that the observed factor structure for the scale does not support their theorised five subscales and so suggest that it is “most parsimonious to conceptualise the ISMIS as measuring a single construct” (p. 16). But when used as a single scale, the ISMIS collapses very different notions. Specifically, the measure fails to distinguish between: the extent to which individuals perceive a mental illness label as self-defining, what individuals perceive a mental illness label means, and what individuals perceive that people without mental illness believe the label means. Instead, this single measure includes highly heterogeneous items assessing experience of received stigma (e.g. “People discriminate against me because I have a mental illness”), endorsement of negative stereotypes about the mentally ill (e.g. “People with mental illness shouldn’t get married”), the perceived applicability of mental illness stereotypes in general to the self (e.g. “Mental illness stereotypes apply to me”), as well as reverse-scored items assessing whether positive stereotypes about the mentally ill apply to the self (e.g. “People with mental illness make important contributions to society”—reverse scored). Consequently, it would be quite

possible that an individual could score highly on this “self-stigma” measure without actually internalising mental illness as being valid and self-defining (for example, by strongly endorsing items relating to experience of received stigma and endorsing negative stereotypes of the mentally ill, without believing these are self-applicable). Conversely, an individual could strongly internalise an illness label but score low on the measure (for example, by strongly endorsing items relating to positive stereotypes about the mentally ill).

The third most common measure of self-stigma reported in the meta-analysis by Livingstone and Boyd (2010), was the 40-item Self Stigma of Mental Illness Scale (SSMIS), with four subscales: stereotype awareness, stereotype agreement, stereotype self-concurrence, and self-esteem decrement (Corrigan, Watson, & Barr, 2006). Here again, the four subscales measure quite distinct constructs, and none of the items captures the extent to which respondents perceive that a mental illness label is self-relevant. The subscale that perhaps most closely taps into self-relevance is the “self-concurrence” subscale, which includes items such as “Because I have a mental illness I am to blame for my problems”. But again, the item is double-barrelled. Respondents could disagree either because they see themselves as mentally ill, but reject that they are to blame for their problems, or because being mentally ill is not a particularly important part of how they see themselves—regardless of their beliefs about their blameworthiness for their problems. This is not to reject the possibility that individuals may display self-deprecating beliefs informed by their category memberships. An extensive social-psychological literature indicates that low status groups can derogate the ingroup and favour the outgroup (Jost, Banaji, & Nosek, 2004). But there is also an extensive literature demonstrating that this is far from an inevitable outcome, and can vary systematically depending on: whether individuals perceive the low status ingroup as self-defining (Thoits, 1985); the norms of the ingroup (McGarty, 2001); changing perceptions of

the legitimacy of status relations (Vaughan, 1978), and; the dimensions of social comparison that are attended to (Camp, Finlay, & Lyons, 2002).

It is perhaps not surprising then, that self-stigma researchers, in reviewing the consequences of mental illness labels, note that under some circumstances, being labelled as mentally ill somehow causes certain individuals to experience feelings of low self-efficacy and withdrawal, for others it has little or no effect, and for yet others it provokes a sense of self-righteous anger and a determination to fight for the rights of individuals with mental illness (Chamberlin, 1978; Deegan, 1990). Corrigan and Watson (2002) describe these divergent responses as the “paradox of self-stigma” (p. 36). But the “paradox” emerges only because of conceptual confusion about what mental illness self-stigma is. The mental illness self-stigma literature wrongly presupposes that being given a label (e.g. “you meet DSM criteria for Major Depressive Disorder”) automatically leads to a person *internalising* that label (e.g. “I *am* a depressive”)—a claim, not supported empirically (Crocker & Major, 1989; Thoits, 1985). Second, the mental illness self-stigma literature supposes that internalising a label is inevitably linked to lower global self-esteem. It is beyond the scope of this review to summarise the extensive literature showing the complex relationships between psychologically internalising a low status category membership, domain-specific self-esteem, global self-esteem, personal and collective self-esteem (Hunter, Platow, Bell, Kypri, & Lewis, 1997). Suffice to say it is more conceptually complex than is typically acknowledged within the mental-illness self-stigma literature, and there is certainly little evidence to suggest that internalising a category membership is consistently associated with low global self-esteem.

Interestingly, Corrigan and Watson (2002) acknowledge that an important moderator of the consequences of self-stigma is whether sufferers identify with sufferers of mental illness as a generic group. They argue that identifying with fellow sufferers may buffer the

negative consequences of received stigma (Crabtree, Haslam, Postmes, & Haslam, 2010).

This argument is consistent with Branscombe, Schmitt, and Harvey's (1999) influential rejection identification model, which demonstrates how low status group members provide a psychological resource for collectively resisting stigmatising beliefs, for example by providing psychological support to fellow group members, and for generating alternative ways of evaluating the group. Importantly, rather than attempting to eliminate stereotyping and categorizing from human cognition, if indeed such a feat were possible (Bruner, 1957; McGarty, 1999), low status group members achieve these ends by changing the meaning of negative stereotypes (Reynolds, Oakes, Haslam, Nolan, & Dolnik, 2000). This suggests, contrary to modified labelling theory, that disorder labels are not straightforwardly and inevitably disempowering, but can provide the basis for a shared identity that enables stereotype change.

But Corrigan and Watson's acknowledgement of these processes points to a further confusion within the self-stigma literature. This literature provides very little theoretical detail about the nature of the "self", other than a tacit assumption that the self is psychologically individuated. But an alternative view, mentioned in earlier chapters, is that group memberships are an integral part of the self (Turner et al., 1987) and so psychologically internalising a category as self-defining is at least partly related to psychologically identifying as a group member (Leach et al., 2008). In the absence of a more specific elaboration of what it means to internalise an illness label, without in some sense identifying with category members as a group, it is very difficult to reconcile the self-stigma literature with Corrigan and Watson's (2002) acknowledgement that identifying with mental illness sufferers buffers the effects of stigma.

A final problem with the self-stigma literature is that, even if it is assumed that individuals invariably internalise mental illness labels, and this is strongly associated with

lower self-esteem and self-efficacy, the inference that this is a causal relationship is dubious. The evidence base for mental illness self-stigma is overwhelmingly correlational (Livingstone & Boyd, 2010), and so the counterfactual is seldom assessed: how does identifying as mentally ill compare to the alternative social roles or identities that are perceived to be available to an individual with mental illness? As Scheff (1974) points out, mental illness labelling occurs when people fail to live up to wider societal expectations and standards of behaviour. Almost by definition, this limits the viability of sustainably occupying alternative social roles. From the perspective of the mental health consumer, therefore, being mentally ill may be far preferable to the perceived alternatives of being criminal or indolent. Indeed this was the very historical impetus for reconceptualising deviancy from being viewed as a moral defect to a medical problem. But again, modified-labelling theorists seem to display a very limited appreciation of the socio-historical context in which the very idea of mental disorder is situated. Consequently, the theory provides few insights into the social-psychological processes that might underlie *both* stability and change in stigma effects.

Self-Stigma and Stereotype Threat

Notwithstanding the conceptual problems within the self-stigma literature, Link and Phelan (2006) also argue that negative consequences of labelling can arise from stereotype threat (Steele & Aronson, 1995). Under stereotype threat, individuals who identify with a relatively low status group, and are then reminded of that membership and asked to perform a task for which group members are stereotyped to perform poorly, can perform worse in the stereotyped domain than when their identity is not salient. For example, African American students, when made aware of their African American group membership, perform worse on tests of intelligence than when their African American identity is not salient (Steele & Aronson, 1995). These effects have been demonstrated in relation to a wide range of

domains, including those diagnosed with depression (Quinn, Kahng, & Crocker, 2004), schizophrenia (Henry, von Hippel, & Shapiro, 2010) and neurological deficits (Kit, Tuoyyo, & Mateer, 2008). Quinn et al. (2004), for instance, demonstrated that individuals with depression perform worse on a test of reasoning ability when their identity as a depression sufferer is made salient compared to when it is not, whereas those with eating disorders do not show similar performance decrements. The authors interpreted this as demonstrating that stereotype threat effects are quite specific to the content of the stereotype. Consistent with this idea, individuals suffering schizophrenia are perceived by naive observers to have worse social skills when the individuals with schizophrenia are lead to believe that the mentally healthy person with whom they are interacting knows they have schizophrenia (Henry et al., 2010).

An influential theoretical explanation for the phenomenon of stereotype threat is that individuals experience a cognitive conflict between a positive evaluation of their self-concept (and an expectation of personal success) and a primed social stereotype of poor performance (Schmader & Johns, 2003). This is thought to trigger an acute physiological stress response, excessive monitoring of performance in the task domain, and an attempt to suppress negative self-evaluations. In turn, these cognitive and affective processes are thought to diminish working memory efficiency (Schmader & Johns, 2003; Shapiro & Neuberg, 2007) which ultimately decreases performance. In this way, labelling can become a self-fulfilling prophecy—at least to the extent that labelling causes stress that impedes performance in a stereotype relevant domain.

The stereotype threat literature appears to present a more robust challenge to the notion that internalising disorder categories as self-defining could in some sense be adaptive. However, again, there is evidence that this process is far from inevitable. Smith and Postmes (2011) provide important empirical evidence showing that individuals can ameliorate the

effects of stereotype threat by challenging what it is that a stereotype means. In the first of two studies, Smith and Postmes asked female participants to either affirm or challenge the stereotype that men are better at maths than women and were asked to do this either individually or in small groups with other women. The authors found that the women subsequently performed best in a test of their individual mathematical abilities in the collective challenge condition and worst in the collective affirmation condition. Importantly, the individual contestation condition did not have the same positive effect on performance. Mediation analysis and content analysis of these data provided important insights into why these differences were observed.

First, the authors' content analysis revealed that, in the collective affirmation condition, female participants tended to theorise and *justify* why women were worse at maths than men and to essentialise these differences by attributing them to biological and evolutionary causes. By contrast, in the individual affirmation condition, participants were significantly less likely to make such theoretical attributions. But conversely, participants in the collective contestation condition were much more likely to mention theories that delegitimised the validity of the stereotype compared to the individual contestation condition. Mediation analysis clarified the mechanism further. The authors demonstrated that, compared to the individual affirmation condition, group affirmation lead to a higher level of perceived consensus about the validity of the stereotype that women are worse than men at maths and this in turn predicted worse performance on the maths test. In a second, similar study, the authors found that confidence in maths ability mediated the relationship between group (vs individual) contestation and maths performance.

The authors interpreted these findings as demonstrating that what makes stereotypes powerful drivers of behaviour is when they are *shared* with fellow members of a group. In this way, dominant stereotypes could be either strengthened or actually reversed in discussion

with fellow group members. Group members therefore provide a resource for assessing the validity of stereotypes and for confirming the nature of social reality (McGarty, Yzerbyt, & Spears, 2002). In turn, it is this socially validated data that guides, and informs behaviour (Turner, 1985).

Although these findings by Smith and Postmes (2011) on gender stereotyping are in a different domain from mental illness stereotyping, they highlight significant theoretical points for understanding the social psychology of stigma in mental illness. The findings underscore the importance of distinguishing stigma, perceiving negative stereotypes held by other groups, and either internalising or rejecting these stereotypes as being self-relevant (Crocker & Major, 1989). The findings also suggest that the process of internalising a stereotype is not a matter of abstracted, individual ratiocination, but is embedded in group processes. Despite the often taken-for-granted assumption that the self is individuated, the findings provide further evidence for the argument that the process of internalising stereotypes as part of the self-concept is a collective achievement.

Summary and implications of Labelling Theory research

Labelling theories and stereotype threat analyses have distinct implications for understanding how being a member of a disorder category might affect mental disorder symptoms. Scheff's labelling theory was an important attempt to explain how self-stereotyping might explain how disordered behaviour gets in under the skin, even if it is unclear how this happens or this changes. Modified labelling theory by contrast, does not emphasise the role of stereotype conformity and the so-called etiological hypothesis (that labelling has a direct causal role in mental illness). Instead, the account supposes that any causal effects of illness labels are largely indirect, secondary consequences of the stress response. As outlined above, it is largely the stress associated with received stigma that exacerbates pre-existing mental illness. In the case of stereotype threat, it is the acute stress

response caused by a sufferer perceiving a “threat in the air” (Steele, 1997 p. 613) that is responsible for decrements in stereotype relevant domains.

While these accounts avoid some of the problems of the vaguely specified mechanisms in Scheff’s labelling theory, they also largely ignore the potential strengths of stereotype conformity as a way of explaining how symptoms co-occur. The mechanisms of stereotype threat and modified labelling theory cannot explain, for example, how particular psychiatric syndromes might emerge in certain cultural contexts, but not others, or the sorts of collective processes alluded to by Hacking (1999), whereby the pattern of illness symptoms dynamically change as a function of interactions between fellow illness sufferers and responses by institutional authorities.

Content of disorder categories: essentialism, malleability and motivation to change

The empirical literature on the consequences of internalising labels points to the pivotal role that beliefs about the nature of mental illness play in shaping behaviour. In this next section I explore these so called folk or lay beliefs about mental illness in much more depth. These beliefs have been the subject of much theorising (Chiu, Hong, & Dweck, 1997; Demoulin, Leyens, & Yzerbyt, 2006; Haslam, 2002a, 2011) and have important implications for understanding stigma, motivation and behaviour as a dynamic process. I assume that two (broad) conceptual dimensions summarise important variation in folk understandings of disorder. These two dimensions relate to (1) the perceived valence of behaviours, and (2) the perceived malleability of behaviours. Although much of the literature reviewed here construes these two dimensions as individual difference variables, I will review evidence suggesting they are in important respects socially shared and contested beliefs. This is important because these two dimensions will play a crucially important role in elaborating the social psychological mechanisms underpinning the Disorder Identity Framework that I explain in Chapter 4.

Lay psychiatric theorising

Haslam (2005) has developed the most well-elaborated analysis of lay psychiatric beliefs. Haslam proposes four dimensions of beliefs about mental illness, each of which is associated with distinct cognitive underpinnings. The latter point is important. Haslam notes that it is simplistic to assume that lay beliefs about psychiatric disorders are simply rough approximations of scientific opinion, as filtered and disseminated through popular media. Instead, Haslam notes that the acquisition and representation of mental illness categories is invariably driven by broader lay theories, cultural understandings, and inherent constraints in cognitive processing. The first dimension proposed by Haslam is pathologising—the lay judgment that a behaviour is infrequent, difficult to explain, and internally attributed. Implicit too in this dimension is that a behaviour is in some sense undesirable. Haslam further notes that this first dimension is separate from the issue of what causes the perceived deviancy. The other three of his four proposed dimensions therefore capture cognitively distinct explanatory modes for understanding why this deviancy occurs.

Haslam's analysis of these three explanatory dimensions draws on Malle's (1999) attribution analysis distinguishing between two kinds of generating factors invoked to explain behaviour: reason explanations (which are intentional) and causal explanations (which are unintentional). This distinction, first made by Heider (1958), has been largely neglected in much attribution research (Kelley, 1967). Classic attribution theories instead focus on the distinction between person and situation causes and have been criticised (Malle, 1999) for failing to explain why lay attributions are typically focused on judgments of intentionality, rather than covariation patterns between behaviours, people and situations. Importantly for the present discussion, lay attributions about intentionality are central to judgments that a particular behaviour is praise or *blameworthy*. Since *ought* implies *can*, blaming somebody because they *ought* to behave differently implies that behaviour is under volitional control.

Specifically, it implies that a behavioural outcome is a result of an actor's intentions (Malle & Knobe, 1997, Malle, 1999)—including the desire to behave a certain way, a belief that certain actions will lead to that behaviour, and a conscious awareness of performing the behaviour. If the behaviour occurs in the absence of these intentions, then the actor could not have behaved otherwise. For example, if behaviours associated with depression, type-1 diabetes and Downes-syndrome are caused by genetic abnormalities, unmediated by a sufferer's intentions, then these behaviours are not blameworthy. Conversely, if gay sex, grandiose delusions, and problem gambling are explained intentionally, and deviate from social norms, then engaging in them opens one to moral reproach—for being perverse, self-centred, or reckless. Therefore, moralising judgments are intimately entwined with folk beliefs about whether deviant behaviour is intentional.

The third lay dimension of folk psychiatry proposed by Haslam is medicalising—roughly the belief that behaviour is caused by some biological essence and is therefore unintentional. More specifically, folk or psychological essentialism (Medin & Ortony, 1989) is the view that deep essences serve to explain or constrain the co-occurrence of surface level properties. Unlike scientific essentialism, critiqued in the previous chapters, folk essentialism does not necessarily purport to be directly informed by scientific evidence (although it may be) and instead involves inferring the existence of an (often vaguely defined) essence placeholder. These place holders may have heuristic value for social actors, even if, as has been argued, their scientific basis is dubious.

The importance of folk essentialism in governing a wide range of related inferences in social perception has now been well documented (Rothbart & Taylor, 1992). Rothbart and Taylor argue that many socially-constructed categories are falsely perceived to be natural kinds and that these judgments are associated with beliefs that the properties of category members are unchangeable and inductively potent. Subsequently, Haslam, Rothschild and

Ernst (2000) demonstrated that these specific essentialist beliefs are embedded in an even wider network of inferences about social groups. Haslam et al. examined lay beliefs regarding 40 different social groups and observed two distinct sets of beliefs: a natural kind dimension—associated with beliefs that the category is historically invariant, immutable, sharply bounded, and naturally occurring; as well as a reification dimension—associated with beliefs that a category is inductively potent, and grounded in deep, inherent properties. In contrast to labelling theories, which tend to treat the category “the mentally ill” quite homogeneously, the data of Haslam et al. (2000) suggested that the two different psychiatric disorders examined, depression and schizophrenia, were understood quite differently. Both disorders were perceived to be moderate on the reification dimension, but whereas schizophrenia was neutral on the naturalness dimension, depression was rated low—suggesting the latter is perceived to be more socially constructed than natural.

In a similar vein, Anh, Flanagan, Marsh and Sanislow (2006) found that although lay people and medical professionals tended to essentialise psychiatric categories to a lesser degree than medical illnesses, lay participants’ views about mental disorder were not devoid of essentialising beliefs, and again this tended to vary depending on the nature of the particular disorder. This included judgments that, to varying degrees, disorders have discrete causal essences, even if these are not yet known, and that these causes needed to be removed for a sufferer to be cured. Finally, Haslam and Ernst (2002) provided experimental evidence that manipulating one essentialist belief lead to changes in other essentialist beliefs. For example, by telling participants that a condition was difficult to cure, participants inferred that the condition was also biologically based, historically invariant, and that disorder sufferers were a uniform group and shared many similarities.

In sum, the findings suggest that folk essentialism links together inter-related inferences about the causes and consequences of mental illness. This network of essentialist

beliefs includes two main, correlated dimensions: beliefs that a disorder is natural and immutable, and beliefs that a disorder category is informative. Again, this body of work suggests these essentialist beliefs are applied differently to different disorders, rather than describing mental disorders as an homogeneous group. Finally, these essentialising beliefs reflect a particular lay understanding about the role of intention. To the extent that a disorder is essentialised it is not under volitional control, and so unlike the folk logic of moralisation, is less likely to be viewed as blameworthy (Weiner, 1985).

Haslam's fourth and final dimension of folk psychiatry is psychologising.

Psychologising is a kind of explanation that involves neither deliberate intentions (as in moralising) nor discrete biological disturbance (as in medicalising). Instead, psychologising is the lay attribution of psychiatric behaviour to mentalistic causes. Often these mentalistic causes are the antecedents of explicit beliefs and desires—what Malle (1999) calls “causal history factors”. Psychologising could include ascribing a mental disorder to the experience of a traumatic event, or personality characteristics, or a troubled childhood. Haslam points out that psychologising is distinct from moralising as a folk explanation, since an individual has no or only partial insight into a psychological cause, whereas reason-based explanations assume the actor is fully aware of his or her intentions, and displays a minimum degree of rationality in integrating them—what Malle (1999) calls the “subjectivity rule” and the “rationality rule” (p. 36). Thus, psychologising is an explanatory mode that acknowledges a grey zone between pure, unintentional causal explanation and fully intentional, reason-based explanation.

Haslam's analysis is an important integration of folk beliefs about mental illness. The analysis highlights how folk psychiatry starts when socially normative explanations fail to explain behaviour. Having determined that some behaviour is deviant, and internally located (pathologising), the question turns to its causal underpinnings. In folk terms: is it in the heart

(moralising), the brain (medicalising) or the mind (psychologising)? By embedding this folk explanatory framework in Malle's attribution analysis, Haslam helps to connect mental illness with its social context. For example, Haslam emphasises that local, social and cultural norms will strongly determine whether behaviours are deviant (pathologising) and how this deviancy is causally interpreted. This may help explain historical and cross-cultural variation in whether and how psychiatric symptoms are perceived. For example, Giosan, Glovsky and Haslam (2001) showed that US citizens tended to perceive a greater number of behaviours as mental disorders, and invoked more psychological attributions for target behaviours compared to Brazilians and Romanians. Glovsky and Haslam (2003) further showed that American acculturation of Brazilians living in the United States was correlated with more psychological attributions for mental disorder. These studies are important in showing how lay understandings of mental illness are socially embedded and how the social meaning of symptoms can vary across time and space.

It is also worth pointing out that Haslam's folk psychiatric framework can be crudely aligned with the two overarching conceptual dimensions mentioned earlier. The question of whether a behaviour is pathological or not can be mapped on to the question of whether some behaviour is negatively or positively valenced. Secondly, the dimensions of medicalising, moralising and psychologising can be roughly mapped on to the dimension of whether a condition is relatively fixed or malleable or in between. Crossing these dimensions opens up the possibility of also describing those behaviours that are understood to be biologically based (fixed) but nevertheless understood as signs of human variation to be celebrated (positive biologising), behaviours that are positively valenced and under intentional control (malleable) and therefore praiseworthy (positive moralising), as well as behaviours that are positive but whose causal basis involves a combination of ability and intentionality (positively psychologising).

Haslam's folk psychiatry framework is important because it illuminates a variety of strategies adopted by people with psychological disorders, and their advocates, to counter received stigma. For example, in recent years, there have been attempts to popularise the idea that mental disorder is an "illness like any other". The US National Alliance for Mental Illness, for example, explains that "just as diabetes is a disorder of the pancreas, mental illnesses are medical conditions" (2015). With Haslam's folk taxonomy in mind, the logic of this is clear. Since moralising is a commonly held stigmatising attitude towards sufferers of at least some psychiatric disorders, promoting a medicalised view of mental illness appears to be a plausible counter-strategy to challenge this belief. If a disorder is medicalised, then it challenges the view that mental disorder is a blameworthy moral failing—caused by deviant desires or weakness of will. Moreover, to the extent that a condition is considered to be uncontrollable, it should engender responses of pity and help (Corrigan, 2000; Weiner, 1985).

The most recent empirical review of this blame reduction hypothesis found that, across 25 studies involving between 4,278 and 23,816 participants, biogenetic attributions for mental illness were indeed consistently and robustly related to lower levels of perceived blame for the condition (Kvaale, Haslam, & Gottdiener, 2013). These data therefore provide important evidence consistent with the lay psychiatric belief model. However, this strategy did not come without costs and also appeared to increase perceptions of dangerousness, a desire for social distance, and, in a related meta-analysis, bio-genetic explanations also tended to induce prognostic pessimism (Kvaale, Gottdiener and Haslam, 2013). Again, there was also some evidence of heterogeneity in these effects—for example biogenetic explanations were associated with blame reduction in the case of schizophrenia, but not depression.

These additional caveats and consequences of medicalising are important for a number of reasons. First, they show that essentialising disorders is no simple panacea for

countering received stigma and may simply open up new avenues for it. Second, the findings highlight a significant empirical limitation in much of this research in that it has overwhelmingly focused on schizophrenia, and to a lesser extent depression. The finding that there is some heterogeneity in the effects of essentialist theories implies that people have at least some awareness of variation in the category “the mentally ill”—and so it seems important to clarify how this variation depends on the nature of the disorder. It seems unlikely, for example, that essentialising social phobia will risk heightening perceptions of dangerousness. But presently there is scant research to address this variation across disorders. Thirdly, the finding that essentialising increases prognostic pessimism suggests an antagonism between perceptions of blame and perceptions of control and optimism for recovery-as-an-outcome. The implication of this is that there is no free lunch in adopting an adaptive strategy from the perspective of sufferers. Conceding that change is possible, that a condition is not fixed, exposes disorder identifiers to the threat of being blamed for their condition. But conversely, while essentialising a disorder is in some ways highly adaptive in guarding against this risk, it comes at the cost of the alternative adaption strategy of attempting to change the self. This trade-off highlights the importance of understanding the motivational consequences of psychiatric lay theories from the perspective of consumers. While Haslam’s folk psychiatry framework has predominantly been used to examine folk beliefs in the general public, rather than consumers specifically, there is a substantial literature examining how self-attributions influence behaviour which can be aligned with Haslam’s work.

Self-theories and achievement motivation

Attributional research over several decades shows how causal understandings of behaviours not only influence observer ratings of targets, they influence actor self-understanding and behaviour. For example, according to Weiner’s attributional theory of

achievement motivation (1985), actors' motivations are directed by the interaction between expectation and value. If some goal is valued, actors engage in a causal search to evaluate whether and how that goal can be attained. Weiner argues that it is highly adaptive if actors can understand the causal structure of the world so they can judge whether goal attainment is constrained by the situation or the properties of the self. Drawing on classic distinctions within attribution theory, Weiner argues that achievement motivation is driven by whether the causal processes underlying goal attainment are internal or external (Heider, 1958; Rotter, 1966), stable or unstable, and controllable or uncontrollable. Attributing failed goal attainment to different combinations on these causal dimensions predicts subsequent motivational states, including helplessness (external, stable, uncontrollable), self-blame (internal, stable, controllable), fatalism (external, unstable), and hope/ renewed effort to attain the goal (internal, unstable, controllable).

Related attributional analyses have been very influential clinically, particularly in the analysis of depression. Individuals with clinical depression tend to feel general hopelessness because they attribute failures to causes that are internal, stable and affect multiple life domains (i.e., that are global, see Abramson, Metalsky, & Alloy, 1989). Individuals who make these attributions tend to make negative inferences about the self (e.g. "I'm worthless") leading to low mood directly, as well as generalised hopelessness, withdrawal, low positive reinforcement, and further low mood. There is also evidence these attributions are causally related to clinical depression (Chan, 2012; Peterson & Seligman, 1984). Conversely, motivation and recovery-as-an-outcome are predicted when individuals attribute failure externally, or to a lack of effort (internal, unstable, controllable).

While these analyses suggest that attributions about the locus, stability and controllability of causes determine achievement motivation, Weiner (1985) emphasises that controllability is the most empirically predictive of these attribution dimensions. This claim

is supported by an extensive body of research by Dweck and colleagues (Dweck, Chiu, & Hong, 1995; Dweck & Elliott-Moskwa, 2010), investigating individually held “self-theories” about whether a personal attribute is relatively fixed or malleable. Individuals with a malleable self-theory are more likely to attribute failure to a lack of effort or strategy and see these experiences as an opportunity to learn. Conversely, those with a more fixed theory tend to attribute failure and success to unchangeable personality characteristics. Consequently, fixed theorists tend to respond defensively: avoiding situations where there is a risk of failure, and rejecting negative feedback that may help them improve.

Dweck’s analysis and the related attribution theories reviewed here highlight how lay theories do not just have consequences for *received stigma* (how stigmatisers attribute responsibility to the stigmatised) but powerfully influence motivations, emotions and behaviours when applied to the self. However, a limitation of these analyses is the relative neglect of how, why and from where these malleability beliefs arise within individuals. Dweck’s initial research suggested, perhaps ironically, that these beliefs are relatively stable, chronically accessible individual difference variables that are reinforced from a young age (Chiu et al., 1997). Similarly, attribution theories suggest attribution styles are often trait-like—for example, Rotter’s locus of control distinguishes between internals and externals, and Abramson et al. (1989) argue that the depressogenic attributional style is an individual difference variable (see also Ickes & Layden, 1978).

A more developed analysis of the origins of these beliefs is outlined in social cognitive theory (Bandura, 1977, 1987). The theory emphasises how self-efficacy, the belief that an individual has the ability to succeed in a specific situation, emerges from one’s social experiences and observational learning. Individual mastery experiences are hypothesised to be the most important determinant of self-efficacy and this has been robustly confirmed empirically (Bandura, 1997). Bandura hypothesizes three further determinants of self-

efficacy: social modelling, persuasion, and physiological arousal. Evidence for these sources is more mixed (see Usher & Pajares, 2008, for a review in the context of educational self-efficacy). Of particular relevance to the present discussion is the evidence regarding social modelling and whether similar others influence self-efficacy beliefs. Usher and Pajares (2008) note that social modelling by observing similar others in an academic context is highly variable, with correlations between vicarious sources and efficacy beliefs ranging from .09 to .56. The authors attributed this variation to the inconsistency in conceptualising and operationalising the nature of “similarity”, leading to between-item reliabilities on measures of similarity that were typically very low (.5 to .6). As Bandura notes, people share myriad similarities and differences and so this poses problems for operationalising what counts as “similarity” in the modelling process.

Despite these ambiguities, Bandura (1968) pioneered experimental studies in a clinical context that provide some important insights into the modelling process. A review of this literature by Schunk (1987) found across several studies (Klorman et al., 1980; Vernon, 1974, Kornhaeber & Schroeder, 1975) that sufferers of various anxiety disorders who watched models displaying coping responses towards an anxiety provoking task (e.g. who verbalised their anxiety, made some mistakes, but worked persistently towards success) subsequently showed less anxiety and performed better compared to watching models who displayed mastery responses (responded confidently and without anxiety). In one study, girls with a demonstrated phobia of snakes showed the biggest reductions in snake phobia after they had watched a coping peer touch a snake, compared to when they had watched peers with mastery responses, or adults displaying either response. Although the authors did not measure self-efficacy directly, they noted that the girls appeared to be particularly influenced by peers who were similar and in a way that was relevant to the task at hand. “Relevant” here is perhaps as question begging as “similar”, but a plausible understanding of this is that

participants saw the anxiety displayed by the coping peers mirroring the anxiety that observers felt themselves. These peers then, seemed to be more informative about the possibilities for the self. This highlights an important point that identifying with fellow sufferers may provide a resource for testing the viability of attempting to change the self. If similar others can change, this suggests that either I too can change or that these others are not really like me in a relevant way. The issue, then, hinges on developing a more conceptually robust account of what “like me” means.

In sum, the achievement motivation, lay theorising, and self-efficacy literatures are critically important in showing how motivational processes are governed by value and expectancy—which again more or less correspond to the basic dimensions of valence and malleability. These theories provide robust evidence demonstrating how causal theories, in particular, play such a crucial role in governing avoidance and approach behaviours, and motivation in the face of set-backs. But the theories provide a relatively under-developed analysis of the social origins of these lay theories. The arguable exception to this is the case of social cognitive theory, which at least examines the role of social modelling in determining self-efficacy beliefs. However, without a more developed theoretical account of social influence, the account struggles to empirically operationalise how this modelling process works.

Attribution and motivation as a dynamic, group process

The foregoing summary of the attribution and social cognitive literature also risks obscuring the social dynamics in which these attributional processes occur. For example, it would be a mistake to assume that beliefs about self-efficacy, and the malleability of attributes, can be neatly partitioned into implications for achievement motivation and implications for perceptions of responsibility, a point acknowledged by many (Weiner, 1985; Haslam, 2005). But this impression can be reinforced by the relative neglect of group

processes in much of this literature on achievement motivation and self-theories. Perceptions about the malleability of attributes are theorised to be largely determined by an individual's personal history of mastery. To the extent that the social context impinges on malleability beliefs, it is through ambiguously specified vicarious influence processes, or through clinical interventions aiming to teach people to make more adaptive attributions (Blackwell, Trzesniewski, & Dweck, 2007). While these analyses acknowledge that motivation is a function of value and expectancy, they invariably hold constant the value dimension in order to tease out the determinants of expectancy/malleability. While this is understandable as a methodological strategy, it perhaps gives the impression that beliefs about the malleability and value of attributes emerge in a social vacuum.

A more dynamic picture emerges when beliefs about the possibility and desirability of changing attributes are assumed to be socially embedded and contested. As demonstrated by Haslam (2005), perceptions of controllability are often at the heart of mental illness stigma and, among other things, govern the viability of the sick role. Indeed discourse about mental illness is often saturated with competing claims about not just controllability but also tightly related folk beliefs about essentialism, discreteness, informativeness and biological determinism. Moreover, to the extent goal attainment is uncontrollable, people often make a virtue of necessity by actively and collectively reinterpreting the valence of the "goal" or constructing positive meaning on alternative dimensions of comparison (Lemaine, 1974). This is a group process, since it involves challenging wider social and institutional beliefs about the meaning of "symptom" and "illness".

There are several lines of evidence for this more dynamic picture. Perhaps most striking are analyses of the historical trajectories of deviancy that were mentioned in the previous chapter in relation to autism and homosexuality (King, 2003). To varying degrees, lay beliefs about both of these conditions have included moralising, medicalising,

psychologising, and depathologising. Ironically, it was the labelling, moralising, and medicalising that created these conditions as a subjectively meaningful way of being and which ultimately lead to collective advocacy movements seeking to depathologise these conditions (King, 2003). The point is not to claim these trajectories are inevitable, but rather that social groups were responsible for contesting and changing lay beliefs about the boundaries and causes of these forms of putative pathology.

An experimental study by Morton and Postmes (2009) provides quantitative evidence for this argument that group dynamics shape essentialist beliefs. The study examined how different kinds of received stigma from majority groups influenced patterns of essentialising by minority group members. Morton and Postmes showed that highly identified gay and lesbian participants tended to endorse the biological basis of homosexuality when asked to reflect on times when the heterosexual majority ignored or neglected to acknowledge homosexuals. However, when asked to reflect on instances where the majority treated them negatively because of who they are, they tended to downplay the significance of the biological basis of homosexuality. This pattern did not emerge among gays and lesbians who did not strongly identify as such.

These findings illustrate two main points. Firstly, the fact that these effects emerged only for high identifiers again shows the importance of the subjective meaning of homosexuality for the self—in Hacking’s (1999) terms, these highly identified group members were not “indifferent” to homosexuality as a kind. Secondly, the study provides empirical support for the claim that from the perspective of the deviant group, essentialist beliefs are not static representations—they can be deployed dynamically and strategically to manage the relationship with the high status group. Participants in this study were sensitive to the different implications of essentialism in different contexts and adapted their beliefs accordingly. As emphasised in Haslam’s folk psychiatry framework, essentialism can be

counter-productive in feeding in to negative stereotypes that deviant groups are fundamentally different from the high status group. But conversely, strategic essentialism helps counter received stigma that deviants are to blame for their condition or, as shown by Morton and Postmes's (2009) research, that a condition is merely a lifestyle choice.

Quantitative social-psychological evidence for these more dynamic processes in the context of mental illness is scant. However, Dumit (2006) provides interesting qualitative evidence for minority group essentialism in the context of contested illnesses. Radical advocates for chronic fatigue syndrome, for example, seek to biomedicalise their condition. The movement actively resists depathologising or psychologising, with some advocates even resorting to death threats against researchers who find evidence for the efficacy of psychological treatments, because of the implication the condition is somehow less real (White, Chalder, & Sharpe, 2014). Again, these qualitative findings highlight how essentialist strategies serve to legitimise claims to occupy the sick role (Parsons, 1951).

This line of research also suggests that the valence of symptoms is contested. Clarke and James (2003), for example, found in qualitative research that some chronic fatigue sufferers in their study reported not wanting to get better if that meant returning to their previous life—and had actually come to value their new sense of self as a disorder sufferer. Significantly, they particularly valued being a part of a larger, more meaningful group. Whatever one's judgments about the merits of this as a way of life, being a chronic fatigue sufferer, and being part of a collective that seeks to resist dominant medical opinion, was actually a source of positive esteem (see also Corrigan & Watson, 2002). As Gamson notes, "there is a big difference between individual resistance, and 'being part of the Resistance'." (Gamson in Taylor, 1996, p. 15; see Dumit, 2006).

Summary and implications

In this chapter I reviewed empirical evidence examining: 1) the causes and consequences of internalising disorder related categories, and 2) how the meaning of disorder categories changes attitudes, beliefs and behaviours. In relation to the first issue, I summarised how labelling theory and modified labelling theory have been influential accounts of how disorder labels are internalised. Both theories suppose that consumers are largely victims of mental illness labels. I have argued that while labelling theories seem to pose a grave challenge for the claim that Disorder Identities can in some sense be adaptive, closer inspection reveals how theoretically and empirically problematic much of this research has been.

Labelling theory suggests that individuals who experience life problems are often labelled as mentally ill and conform to wider societal stereotypes about mental illness, partly because when in a distressed state they are highly suggestible and partly because institutional authorities reinforce conformity to the mentally ill role. But the precise mechanisms underlying this conformity are unclear and so the theory cannot explain how a generic stereotype of “being mentally ill” generates specific symptom constellations, nor how these stereotypes can systematically change.

By contrast, modified labelling theory proposes that mental health consumers are victims of labelling in a different way. Modified labelling theory suggests that labels do not directly cause mental illness symptoms but rather cause stigma which exacerbates prior individual vulnerabilities. Modified labelling theory has made an important contribution to the literature in highlighting the pervasiveness and negative consequences of structural and societal discrimination towards those with mental illness. But the modified-labelling analysis does not consistently distinguish between being labelled, perceiving a label as self-defining, and the meaning of a label. Consequently, the theory supposes that mental illness labels

cause individuals to automatically and inevitably internalise negative societal stereotypes about mental illness. However, the theory has great difficulties in explaining empirically how, in some circumstances, labels are associated with low self-esteem, but in other circumstances are empowering (hence, “the paradox of self-stigma”).

In the second section of this chapter I reviewed evidence showing how the meaning of disorder categories can have divergent consequences for mental health consumers. Two important content dimensions emerged from this analysis: 1) the valence of disorder “symptoms”; and 2) the malleability of these symptoms. The valence dimension in some ways problematises how one even refers to disorder. Valence governs whether behavioural deviance is a disorder to be endured, or perhaps ameliorated, or legitimate human variation to be accepted or perhaps celebrated. Moreover, the valence of deviant behaviour has motivational consequences for deviants—since motivation to change is a function of value (whether some behaviour is positively valenced or not) and expectancy (whether a person perceives that a behaviour can change).

This leads to the second dimension of lay beliefs about disorder categories: the causal basis of deviancy. As discussed, this has been further refined in different ways across different attribution literatures, such as locus, stability, and controllability (Weiner, 1985), malleability (Dweck, 1995), and the intentional and non-intentional (Malle, 1999). Haslam’s folk psychiatry framework embeds these lay understandings in a broader social context and highlights the links between causal attributions and moralising, medicalising, and psychologising.

However I have argued that, with important exceptions, these social psychological approaches have understood these dimensions as relatively stable individual differences, largely shaped by the accumulation of individuated experience and abstracted from its social context. In contrast to this view, I have emphasised throughout this chapter the empirical

evidence that shows how these beliefs play out in a contested and dynamic inter-group context. This evidence aligns much more strongly with the argument that mental health consumers are not necessarily passive recipients of diagnostic labels and nor do they necessarily endorse dominant societal or lay understandings about what these labels mean.

The examples touched on here also suggest that fellow sufferers are often an important resource in reaching consensus about what a disorder means in a social context. They help the consumer formulate whether, for example, changing and overcoming one's condition is possible and desirable; or that a condition is a biomedical problem, beyond intentional control; or that the "symptoms" of deviancy are in fact legitimate ways of being. Moreover, these views can often be threatened by both orthodox medical opinion and stigmatising lay beliefs from "normals". The challenge, then, is to develop an alternative social-psychological account that illuminates these intergroup dynamics, while at the same time accounting for, and building on, the substantial research showing the behavioural and motivational consequences of lay theories.

Chapter 4: A Social Identity Analysis of Disorder and Recovery

The aim of the current chapter is to finally elaborate the social identity based mechanisms that could explain how Disorder Identities systematically change the properties of HPC disorder categories. In doing so, this chapter brings together the HPC and Disorder Identity Frameworks developed in Chapters 1 and 2, and the discussion of lay theorising about disorder from Chapter 3, to show how mental disorder and recovery can be analysed empirically as a group based process. While I have made this claim at various points in the discussion so far, and have referred to the Social Identity Approach, I have not yet *demonstrated* the social identity based processes that could link Disorder Identities to the dynamics of HPC categories.

As I have alluded to at various points in the discussion, the Social Identity Approach is fundamentally anti-individualistic in the sense that it assumes that psychological processes dynamically interact with the social environment (Turner & Oakes, 1986). The theory therefore attempts to explain uniformities in social behaviour not as a function of intra-individual mechanisms and inter-personal processes, but as an outcome of the interaction between individuals, their self-representation as group members, and the social world. In other words, it assumes that individuals and groups mutually and interactively determine each other. This contrasts with the dominant, individualistic meta-theory that underpins the biomedical model, outlined in Chapter 1, and many of the cognitive theories outlined in Chapter 3 (most notably modified labelling theory, and social cognitive theory). The latter approaches suppose that “group behaviour” is largely reducible to individual and interpersonal behaviour.

In the first part of this chapter I outline in detail the core tenets of the Social Identity Approach. The second part of the chapter explains how these tenets can be applied to explain how Disorder Identities could enable group members to adaptively modify themselves and

their social environment. Specifically, my analysis allows for systematic predictions about how shared beliefs about the nature of the group and the socio-structural context will determine the adaptive strategies that disorder identifiers adopt. In turn, the consequences of these group-based strategies can be mapped on to the HPC-based framework developed in Chapter 2. This starts to reveal what Disorder Identities mean from the frame of reference of identifiers while at the same time showing the homeostatic, homeodynamic, heterostatic, and heterodynamic consequences of Disorder Identity Adaption Strategies within the HPC framework (from the frame of reference of experts). By mapping these mechanisms more systematically, it is possible to see the broader Disorder Identity Analysis as a truly dynamic system. Moreover, because of the flexibility of the underlying self-categorization mechanisms, the morphology of particular disorders can be predicted to display either stability or change, depending on the context. This change does not occur because of mysterious, synchronous changes in individuals, but because those who identify as disordered (Disorder Identifiers) respond collectively to changing threats and opportunities within the environment, informed by their shared understanding of what disorder categories *mean*.

I conclude the chapter by presenting a taxonomy of different Disorder Identity Adaption Strategies, the consequences of these strategies when viewed from the HPC perspective, and the Social Identity based mechanisms that underpin these different strategies. Where possible, I present empirical evidence and examples from the literature of each of these strategies in action.

The Social Identity Approach

Social Identity Theory and intergroup behaviour

SIT is a theory of intergroup relations that seeks to explain differences in the way people respond to being a part of a group-based social hierarchy. The theory posits that intergroup behaviours occur partly as a function of how the socio-structural context is

perceived. Key socio-structural variables include group members' perceptions about the legitimacy and stability of intergroup status differences, and the perceived permeability of the boundary between a low and high status group. The theory further claims that the form of intergroup behaviour is determined by the salience of one's social identity. The construct of social identity was developed by Tajfel (1978) and Turner (1975, Tajfel & Turner, 1979) and is defined as "those aspects of a person's self-concept based upon their group memberships together with their emotional, evaluative and other psychological correlates" (Turner & Oakes, 1986, p. 240). This definition underscores an important distinction between sociological groups (defined in the third person by objectively observed characteristics) and social-psychological groups—groups that are psychologically meaningful to the self (Turner, 1999). This distinction is crucial. Individuals who meet diagnostic criteria for Major Depression are part of the sociological group of Major Depression sufferers. But meeting these criteria does not imply this category is subjectively meaningful—individuals may not see themselves as "a depressive" (part of a psychological group). In the context of the current discussion, it is this distinction that makes it possible to think of disorder from the frame of reference of identifiers (and whether an individual subjectively reports identifying as disordered), and disorder from the frame of reference of experts (in terms of whether a person meets clinical criteria for an HPC-defined psychiatric category). As mentioned in Chapter 2, this further implies four possibilities. Individuals can: identify as disordered and be clinically diagnosable as disordered, not identify as disordered but be clinically diagnosable as disordered, can identify as disordered but not be clinically diagnosable, and can neither identify, nor be diagnosable, as disordered.

In a given context, a social identity may be more or less psychologically active, or salient. SIT originally posited that social behaviour occurs on a continuum from, at one extreme, purely interpersonal behaviour (determined solely by individual characteristics) to, at the other extreme, purely intergroup behaviour (determined only by social group membership). SCT later explained the cognitive processes that underpin this behavioural continuum by elaborating how particular self-categories become salient as a function of features of the social context and the perceiver's background knowledge and motivations. This will be detailed shortly, but for now the point is that in different contexts people can think and act more as individuals or more as a member of a social group.

SIT further claims that generally individuals "strive to achieve or to maintain positive social identity" (Tajfel & Turner, 1979, p. 40). By and large, a positive social identity is achieved through favourable comparisons with other, contextually relevant, groups on some dimension of comparison. Status, therefore, does not necessarily depend on objective discrepancies in wealth or power, but instead is an outcome of intergroup comparisons.

However, the motivation to achieve or maintain a positive social identity is sometimes wrongly assumed to be the main or the only driver of intergroup behaviour. Rather, it is but one of "three legs of a conceptual tripod" (Tajfel, 1979, p. 185) that interacts with the perceived socio-structural context, and the salience of a given social identity. SIT elaborates how these three elements interactively predict how group members tend to respond to being in a low status group. To the extent that group members perceive their social identity to be unsatisfactory, group members either seek to leave the group, by joining a higher status group, or to make the ingroup more positively distinct. This plays out through three broad social identity management strategies. The first strategy is *individual upward mobility*. This is predicted to occur when boundaries between the low and high status group are perceived to be permeable, prompting individuals to disidentify with the low status group and attempt to

join the higher status group. A point that will be discussed in detail in the next chapter is how permeability is conceptualised in the case of the boundary between illness and health and to what extent upward mobility is necessarily individualistic and individuating. For now, boundary permeability can be understood as whether “it is *possible* (through talent, hard work, or whatever other means) to move individually into another group” (Tajfel & Turner; 1979, p. 35).

Not surprisingly, individual upward mobility is not always possible because intergroup boundaries are often either perceived to be, or simply are, *impermeable*. Under these circumstances, group members have no choice but to find other means to construct a positive and distinct social identity. One strategy is to attempt to positively define the group in other ways: for example, by positively defining the group on alternative dimensions of comparison (e.g. sad but wise, poor but happy), redefining the valence of the status defining dimension (“black is beautiful”), or comparing the ingroup with even lower status groups (as in “poor white racism”, see Leach & Spears, 2008). These group-based social creativity strategies are predicted to occur only to the extent that status relations between the high and low status group are perceived to be legitimate and stable. If, however, the high status group’s authority is perceived to be unstable or illegitimate then the theory predicts that group members will increasingly attempt to directly challenge the high status group and engage in social competition. At the extreme, this strategy involves overt conflict, for example to gain access to resources, and ingroup favouritism on a dimension of comparison valued by both groups.

Empirical support for these core predictions within SIT has been robustly demonstrated. A meta-analysis by Bettencourt et al. (2001) found evidence that the perceived legitimacy and stability of intergroup relations moderated ingroup favouritism (i.e. social competition) and that boundary permeability was associated with decreased

identification with the low status group and a desire for upward mobility (Ellemers, van Knippenberg, De Vries, & Wilke, 1988; Ellemers, Van Knippenberg, & Wilke, 1990; Jackson, Sullivan, Harnish, & Hodge, 1996). In recent years, there has also been some evidence that these kinds of identity management processes operate in clinical contexts (Branscombe, Fernandez, Gomez, & Cronin, 2012; Jetten, Haslam, & Haslam, 2011). Fernandez et al. (2012), for example, examined the experience of individuals with achondroplasia (dwarfism) in Spain, where dwarfism is a highly stigmatised condition, and compared this with individuals with dwarfism in the United States, where there is an active support community called “little people of America”. Fernandez et al. found that, depending on the social context, a strategy of individual upward mobility (by undergoing extremely painful limb lengthening surgical procedures), or a strategy of social creativity, were both effective in improving wellbeing. In Spain, surgery (individual mobility) was effective and one of the only means for the individual to positively define him or herself. In the US, by contrast, a social creativity strategy was effective in contributing to wellbeing because the existence of a social group of fellow little-people imbued this category with positive and distinct social meaning and buffered individuals from the negative consequences of received discrimination.

While the social identity management strategies are sometimes presented as neatly and categorically distinct, Tajfel and Turner (1979) emphasise repeatedly that intergroup behaviour occurs on a *continuum* between acting as an individual and acting as a group member. This behavioural continuum parallels an ideological continuum between belief in social mobility and belief in social change. To the extent that a social mobility belief system is endorsed more widely within a culture, then behaviour is more likely to be individuated (as in individual mobility), whereas to the extent that social change is endorsed then behaviour is more likely to be group based (as in social creativity or social competition). This emphasis

on a behavioural continuum suggests that some identity management strategies are not strictly incompatible. For example, Tajfel and Turner note that where social creativity involves re-evaluating the group on alternative dimensions of comparison, this can create challenges as group members seek to legitimise the new dimension—initially within the ingroup and then to other groups. To the extent that the outgroup is threatened by this, intergroup behaviour may escalate into social competition. In other words, in a particular intergroup context, strategies may include elements of both social creativity and social competition.

Self-Categorization Theory

SCT provides a more detailed account of the cognitive underpinnings of the continuum between individual and group-level behaviour postulated within SIT. SCT does not replace or conflict with SIT, but rather deepens the analysis of how, and under what circumstances, people's actions are determined by their psychological group memberships. A starting point for the theory is that self-perception is hierarchically organised: people can see themselves not just as unique individuals, compared to other individuals, but as members of groups, compared to other social groups. These groups are higher-order categories with varying levels of inclusiveness and which can be nested (depressive, mentally ill) or cross cutting (depressive, poet). The psychological shift in self-perception between perceiving oneself as an individual and perceiving oneself as a category member is marked linguistically by the shift between saying "I" and "we". The theory assumes that this is not a linguistic artefact, but reflects a substantive cognitive change in how one sees oneself in a situation, and which in turn governs one's beliefs, emotions and behaviours. When people self-categorize as group members their perception becomes *depersonalised* and their self is perceived as being more or less interchangeable with other category members. Only under these conditions do the norms, beliefs, attitudes and typical behaviours of fellow group members

become a “guiding image for action”, in Scheff’s (1974) terms. Importantly for the present discussion, it is this depersonalised self-perception that leads to perceived in-group homogeneity.

The cognitive mechanism that makes a particular self-categorization salient is partly a function of individual background knowledge and motivations, and partly a function of the “fit” between some categorization and the immediate social context. The latter is determined partly by the principle of meta-contrast, according to which a particular self-categorization is likely to be salient to the extent that the inter-category differences are perceived to be greater than intra-category differences (the meta-contrast ratio) in a given situation. And so, for example, a group of social and neuro-psychologists in conversation are likely to self-categorize in terms of their sub-disciplinary self-categories. They are also likely to actively accentuate those beliefs and attitudes that differentiate their areas of expertise (perhaps by, respectively, minimalising and emphasising their belief in the importance of biological processes in determining behaviour, see Haslam & Turner, 1992). In other words, the meaning of an identity is not fixed but varies as a function of the comparative context.

Moreover, the salience of a particular self-category can change as rapidly as the context shifts (Antaki, Condor, & Levine, 1996). And so, if physicists joined the conversation, the psychologists may recategorize themselves in terms of their shared, super-ordinate category, “we psychologists”, in *contrast* to the physicists. This re-categorization is accompanied by the perception that the formerly differentiated psychologists are more or less equivalent or interchangeable category members. And again, this recategorization at a higher level, also leads to a shift in norms, attitudes, beliefs and behaviours so that the psychologists are now more likely think and act in terms of the higher-order stereotype “we psychologists” *in contrast to* the physicists (Oakes et al., 1991).

This example might give the false impression that fellow category members must be physically present to trigger group-level categorization. On the contrary, these group-based self-category representations are often implied by the context. For this reason, as I will show later, often dyadic interactions can be understood psychologically as intergroup behaviour. Indeed there is evidence that the physical presence of other ingroup members may, in some circumstances, attenuate the salience of a shared identity (Reicher, Spears, & Postmes, 1995). This is because observable differences (gender, accent, dress, race) may trigger cross-cutting self-categorizations that reduce the comparative fit of an alternate categorization scheme. For this reason, computer mediated communication—for instance in online chat forums where there are relatively few cross-cutting identity cues—can make particular online forum-related self-categories highly salient (Postmes, Spears, & Lea, 1998; Spears, Postmes, Lea, & Wolbert, 2002).

But comparative fit is just one of the constraints on category salience and category meaning. Other determinants include whether the social stimuli fit with the perceiver's normative expectations about how category members are expected to think and behave, also known as the "normative fit" of a stimulus array (Oakes et al., 1991). In other words, the categorization process is not just dictated by the perceived similarities and differences between stimuli. A related point has been made in relation to categorization more generally: similarity alone is an insufficient constraint on categorization, since stimuli can be similar or different on countless dimensions, depending on which dimensions are selected (Murphy & Medin, 1985). And so, background *theories* about the nature of a category help determine whether or not a dimension of comparison is relevant. If the discussion between social and neuro-psychologists shifts to discussing the football, then the relevant supporter identities are likely to be more theoretically relevant / have a better normative fit, than disciplinary

identities, in explaining the observable behaviours and expressed beliefs of individuals in the stimulus array.

Finally, self-categorization is constrained not just by the comparative and normative fit of a situation but by perceiver readiness. Perceiver readiness is determined in part by an individual's long term experience of using a particular social category. This provides a theoretical basis for distinguishing between category salience and social identification (McGarty & Grace, 1999). Social identification is both an outcome and a determinant of category salience. To the extent that some way of categorizing the self has proven to be useful or meaningful in the past, the more readily an individual will be to access and use it in novel situations (of course, as always, constrained also by the other determinants of fit). Moreover, if this category generates favourable intergroup comparisons then the category can be expected to be associated with positive emotional and evaluative consequences that underlie the notion of social identification, as outlined in Social Identity Theory (Tajfel & Turner, 1979). SCT finally postulates that perceiver readiness is further determined by an individual's long term knowledge as well as their current motivations and perceived situational threats.

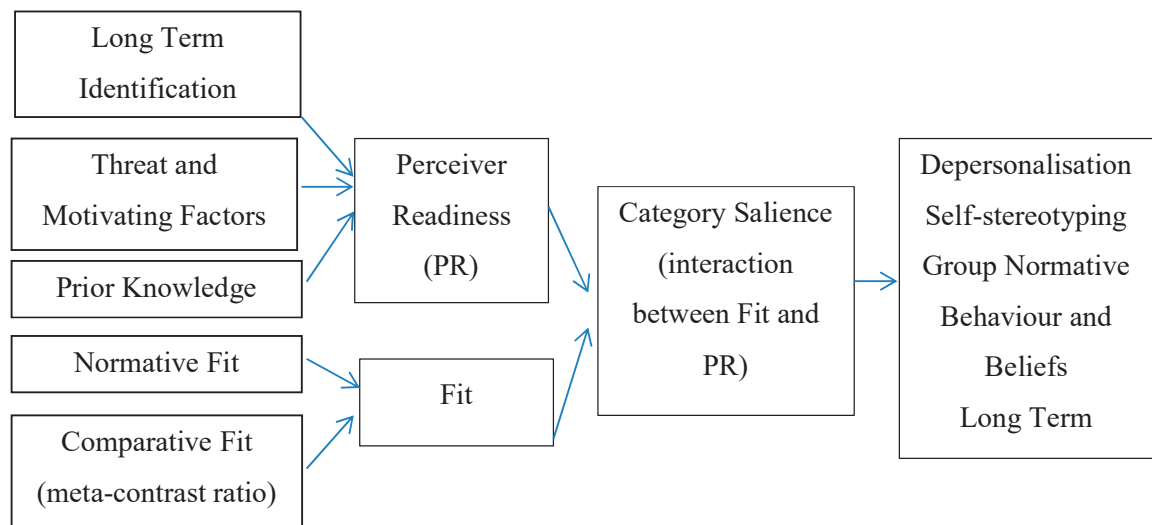


Figure 7: Interaction between perceiver readiness and fit determines category salience and its consequences.

(Note that Long Term Identification is both a determinant and outcome of category salience.)

In sum, SCT posits that self-categories become psychologically prepotent owing to the interaction between the “fit” of a situation and an individual’s background knowledge and experience as an actor in the social world. The determinants of category salience are shown in Figure 7. The theory helps explain how and when psychological group memberships get in under the skin and guide beliefs, attitudes and behaviour. Further, the theory paints a highly fluid understanding of the self. To the extent that beliefs and behaviours are relatively stable, then this is seen as an outcome of the relative stability of intergroup relations and the intergroup context. But, moment-to-moment, the mechanisms outlined within the theory allow for considerable flexibility: by not only governing which self-category is salient, but determining what a category means.

Finally, the theory implies that the categorization process involves the formation of *meaningful* representations, not just of other groups, but of the ingroup itself (what are sometimes called “self-stereotypes”). As with categorization more generally, stereotyping of the ingroup (i.e. self-stereotyping) and the outgroup is understood within SCT as a meaning-

making process—a process that serves to predict, explain and control who we are (and who we are not) in a particular context (Oakes, Haslam, & Turner, 1994).

However, what is perhaps not yet clear is how these representations of the ingroup and the outgroup become socially shared. This is a crucially important point if the self-categorization process is to provide a basis for explaining not just how individual group members perceive and represent similarities across group members, but actually start to think and act in more similar ways. The contrast with categorizing objects is informative. Blades of grass may be categorized similarly because they are relevantly similar in contrast to trees and shrubs. But categorizing these objects does not change the *actual* properties of blades of grass since, in Hacking's (1999) terms, these are "indifferent kinds". By contrast, the self-categorization process leads to a socially shared representation of the self that, according to SCT, leads group members to start to think, feel and act in unison and thereby become more similar. The process by which group members develop a consensus about this shared representation is elaborated in the Social Identity analysis of social influence.

Social identity analysis of social influence

The social identity analysis of social influence (Turner, 1991) challenges early dual process models (Deutsch & Gerard, 1955) within social psychology that distinguished between informational and normative influence. According to the latter perspective, individuals are influenced either because they need information about objective reality (informational influence) or because they seek group affiliation (normative influence). In the case of normative influence, individuals are motivated to publicly comply with social norms, but this "mere compliance" does not translate into deeper, private acceptance. Deutsch and Gerard contrast this with informational influence, which they argue is driven by the need for valid information, and is predicted to occur when individuals are uncertain about the nature of reality and when some information source is perceived as having some objective authority

on the matter. In contrast to normative influence, this informational influence is likely to lead to long-term internalisation of these beliefs.

The social identity analysis, by contrast, rejects this distinction and argues that a shared social category membership provides the basis for social influence, what Turner, Wetherell and Hogg (1989) call referent informational influence. Self-categorization not only accentuates perceived similarities between group members, it leads group members to *expect* to share relevant beliefs, attitudes and behaviours. To the extent these expectations are fulfilled, group members will experience subjective validity about the situation and so relevant beliefs, attitudes and behaviours will be perceived to be valid. By contrast, to the extent that group members *disagree*, or members behave counter-normatively, then this generates subjective uncertainty. This absence of consensus implies that at least one person is wrong—a situation that demands resolution if a stimulus is to be rendered meaningful. This is especially likely if situational cues make a shared identity salient, group members are responding to the same stimuli, and resolving any uncertainty is important. A lack of consensus can be resolved either by group members changing their own perception of the situation (i.e. they are socially influenced) or by searching for alternative ways of categorizing the self and others within the group (e.g., “the others are wrong”) or by reinterpreting the situation in such a way that explains why group members have divergent views. Group discussion can therefore play a critical role in the social influence process and helps generate consensus about who “we are” and “what we believe”. To the extent this consensus is not achieved then it implies that a shared self-category provides a poor “fit” to the situation and alternative categorizations can be tested.

This social influence analysis has further implications for understanding expert and minority influence. Experts, such as doctors in the context of mental health recovery, are likely to be perceived as valid sources of information to the extent they are seen as being

representative of wider, shared, social institutions. This can occur even if at an inter-personal level, doctors and patients occupy very different roles and power relations: doctors can be profoundly influential if they are seen as prototypical leaders of a wider social group one identifies with and that is associated with shared norms about the legitimacy and value of modern bio-medical science, or a particular specialty (psychiatry, psychology) (Turner, 1991; Tarrant, Hagger & Farrow, 2011).

But as will be discussed shortly, often minority group members seek to challenge majority social norms about the nature of reality. Here again the social identity analysis of social influence helps to explain how. It might be thought that, by definition, this minority influence must involve influence across group boundaries, and so violates the principle that influence is an outcome of ingroup consensus. But instead, the flexibility of the categorization process means that effective social influence only occurs to the extent that the minority group can effectively appeal to shared values, within a higher-order identity (David & Turner, 1996).

If minority groups are effective, they present a coherent set of norms that are compatible with aspects of the superordinate identity while generating subjective uncertainty within majority group members about the validity of the status quo. If the influence attempt is too confrontational, and escalates into intergroup competition, then this will lead to group polarisation and ultimately damage the interests of the minority. But if the influence attempt is framed in terms of alternative norms which are compatible with a shared overarching identity, then the minority is more likely to bring about conformity towards a new consensus about who “we” (the superordinate group) are and what “we” believe. The social identity analysis therefore helps to set the stage by explaining not only how it is that group members come to reach consensus about who “we are” in a social context, but how it is that group members can use social influence to modify the social environment.

Applying the Social Identity Approach to understand the dynamics of Disorder Identities

The Social Identity Approach provides a unique perspective on a range of the theoretical issues that have arisen in the discussion so far, including: identity salience, stereotyping, social influence, attribution, and essentialism. In this section, I show how Disorder Identities, when viewed through the lens of the Social Identity Approach, provide a means for group members to modify themselves and the world. I will further aim to document these adaption strategies systematically and demonstrate the circumstances in which they are likely to have homeostatic, homeodynamic, heterodynamic, and homeodynamic consequences, depending on the norms of the group, and in interaction with the socio-structural context.

To briefly reiterate, in Chapter 2 I claimed that Disorder Identities arise when a group of people hold the shared beliefs: 1) that group members currently, or in the past, display(ed) a coherent set of behaviours, feelings, and beliefs; 2) have a shared awareness that these behaviours are, or should be, perceived institutionally as dysfunctional (even if group members reject this); and 3) perceive that the group is a subjectively meaningful way of perceiving the self. I also claimed that Disorder Identity Adaption Strategies occur where: 1) group members perceive that a Disorder Identity is self-relevant in a particular situation, and 2) group members act in a way that is informed by a *shared* understanding of what a disorder category means, and the socio-structural constraints of their environment.

The Social Identity Approach provides a much deeper analysis of these claims. Specifically, SCT explains how and when Disorder Identities become salient. Like labelling theory (Scheff, 1974), the approach suggests that stereotypes about a category can indeed determine an individual's beliefs, behaviours, and emotions. But the approach suggests that self-stereotypes are much more contextually sensitive than is envisioned within labelling

theory. Seeing oneself as an anxiety sufferer need not be an omnipresent self-category—just one of many ways of seeing the self, alongside being a mum, an IT consultant and a cricket tragic. Only moment-to-moment contextual cues govern whether and to what extent seeing oneself as an anxiety sufferer usefully explains a situation. To the extent this self-category repeatedly “fits” as an explanation of the self in context then an individual may come to form a more enduring representation of the self as an anxiety sufferer (i.e. a Disorder Identity). Moreover, to the extent one develops a long-term identification with a disorder category, as a component of perceiver readiness, this increases the likelihood that such a category becomes salient in subsequent interactions.

This identity salience mechanism therefore explains how it is that Disorder Identities can be a disorder-specific homeostatic mechanism, in the sense outlined in classic HPC accounts. HPC accounts require that a “similarity generating mechanism” must exist that explains how the properties of a category co-occur. According to SCT, to the extent that people self-categorize as members of a particular social group then people will perceive themselves as relatively interchangeable with other category members (i.e. depersonalisation) and will thereby conform to the normative beliefs, feelings and behaviours of fellow group members in that situation. To the extent that a Disorder Identity involves a normative understanding that category members are symptomatic (“we depressives”) then their symptomatic experience will converge towards a shared ingroup stereotype (“we depressives are introspective, irritable, somewhat morbid”). It is depersonalised self-perception, therefore, that leads to ingroup homogeneity (Haslam, Oakes, Turner, & McGarty, 1995) and thereby serves as a homeostatic mechanism.

Of course various other cross-cutting identities may also generate certain patterns of psychiatric symptomatology. For example, making particular age (St Claire & He, 2009) or gender related identities salient has been shown to change symptom experience, and this has

been the focus of much work within the social-identity based “social cure” literature (Jetten, et al., 2011). But these cross-cutting identities are of less relevance in the context of an HPC account of mental disorder—since they cannot clearly explain the specificity of disorder symptoms displayed by the particular group of individuals who meet diagnostic criteria for a disorder. To illustrate, it has been widely noted that women are more likely to experience depression than men, and this has been attributed to many bio-psycho-social causes—among them that symptoms of depression may be a more socially acceptable way of expressing and experiencing distress for women than men (Piccinelli & Wilkinson, 2000). But even assuming this is due to identity based processes, it would seem that identifying as a woman would be a relatively insensitive and non-specific identity-based mechanism for explaining how depressive symptoms co-occur, in the sense required by HPC accounts of disorder.

The more difficult case is where individuals identify with a lay psychiatric category (“we worriers / anxiety sufferers”) that is associated with a broad class (anxiety disorders), rather than a particular diagnostic type (social anxiety disorder). This does not necessarily threaten the current analysis, particularly given the high levels of comorbidity in anxiety disorders, but suggests that some Disorder Identities may be more or less specific and sensitive than others. It is also taken into account within the underlying theory of the Social Identity Approach: depending on the context, the hierarchically nested categories “mental illness sufferer”, “anxiety sufferer”, or “social anxiety sufferer”, may be more or less salient.

For Disorder Identities to serve as homeostatic mechanisms they must not only be reasonably disorder specific, there must also be a shared understanding that the disordered behaviour relates to how typical group members are *now*—not just who they were in the past or who they will be in the future. Identifying as a member of AA is a potentially powerful Disorder Identity (the importance of which will be discussed later), but the normative understanding is that most AA members are not current drinkers (and indeed most would not

meet diagnostic criteria for substance abuse disorder). It does not follow, then, that identifying as an AA member will serve as a homeostatic mechanism for alcohol abuse disorder (i.e. causing category members to display more similar behaviours, informed by a heavy drinking norm). Why? Because there is a shared understanding that drinking is precisely what it is that typical AA members do *not* do.

Disorder Identities, context dependent self-stereotyping and symptom change

SCT also helps explain why the meaning of a disorder category is not a static stereotype but might be expected to *change*—and not just in response to the situational features of a particular context, but over longer periods. The Social Identity Approach therefore generates a markedly different understanding of the nature of stereotyping compared to the theoretical perspectives reviewed in the previous chapter. Labelling theory, for example, posits that stable, societally shared stereotypes about insanity are ingrained and reinforced from a young age and ultimately cause labelled individuals to display conformity to a stereotype of “the mentally ill” (Scheff, 1974). But the two problems with this understanding are that such a generic stereotype cannot clearly explain the specific constellations of symptoms observed in particular disorders, nor how these constellations change.

SCT, by contrast, provides a theoretical account that explains how the properties of a disorder can change as a function of social comparison processes. The interaction between fit and perceiver readiness does not merely determine the salience of a category, it also shapes the way that category members self-stereotype. Levine and Reicher (1996) provide evidence for this process in the context of symptom perception and masculine identity. They asked male rugby players to rate the seriousness of various injuries and were told that their responses would be compared either to a group of women or to a group of “new men”. The researchers predicted and observed that a facial scar is perceived as more serious and

distressing in a context involving women relative to a comparative context relating to new men, but that these differences did not emerge for a concealable knee injury. In other words, the perceived significance and distress associated with a symptom systematically varies as a function of which outgroups are contextually salient.

Contextual variation in the ingroup norm is also demonstrated by social identity analyses of group polarisation, showing how shared group beliefs do not necessarily reflect conformity to average individual beliefs. While there are various theoretical accounts of group polarisation (see Isenberg, 1986 for a review), the SCT analysis suggests that beliefs can converge to become more or less extreme, compared to group members' average initial personal beliefs, depending on the social context (Hogg, Turner, & Davidson, 1990). Hogg et al., for example, showed group members' beliefs converge towards a risky norm when contrasting with a cautious outgroup, towards a more cautious norm when contrasting with a risky outgroup, or towards the pre-test individual mean when contrasting with both risky and cautious outgroups.

Although there is relatively limited direct evidence for these processes in the specific context of Disorder Identities, the theoretical implications of these findings are clear: stability and change in the features of disorders may be partly explained by stability and change in the comparative social context. Because of this, shared Disorder Identities can be potentially powerful homeodynamic mechanisms, by causing the average symptom severity experienced by Disorder Identifiers to either increase or attenuate, in *contrast* to a salient outgroup. As will be elaborated shortly, this understanding helps explain the dynamics of certain contested disorders. For example, the number of alters displayed by people with multiple personality disorder increased by an order of magnitude over the space of a decade (Hacking, 1999). For such contested disorders, outgroup representations of the category often dismiss the legitimacy of the disorder or emphasise its psychological, rather than biological determinants.

Ingroup representations are therefore likely to imbue those properties of the category that are seen to signify some legitimate somatic disturbance (in *contrast* to the outgroup representation) as being particularly inter-subjectively meaningful.

While this suggests a more strategic dimension to symptom perception, these processes do not necessarily imply or require explicit collusion on the part of Disorder Identifiers. Group norms are often tacitly held and can form and change very quickly without explicit discussion (Hogg & Reid, 2006; Burgoon, Buller & Woodall, 1996). Nor does this suggest that Disorder Identifiers can straightforwardly “choose”, as individuals, to reject ingroup beliefs, or that individuals bring symptoms upon themselves and are therefore blameworthy. If things were so simple, then cognitive therapies aiming to alter symptomatogenic cognitions would be quick and easy. Instead, the point is that fellow disorder sufferers provide subjectively meaningful and validated information about which emotional, behavioural and cognitive signs should be ignored, and which signs should be attended to, remembered, interpreted, and responded to (a point elaborated at length in Chapter 1). A Disorder Identifier does not therefore experience unusual cognitive, emotional, and somatic sensations as discrete and isolated phenomena, but rather perceives, *and expects to perceive*, these symptoms as part of a wider, more meaningful category that is shared and validated by similar others.

In summary, depersonalisation and context dependent self-stereotyping, as outlined within SCT, provide a plausible social-psychological basis for an externalist homeostatic mechanism and a homeodynamic mechanism respectively. These two mechanisms are likely to operate only to the extent that:

1. the situational context makes a Disorder Identity a salient and meaningful way of defining the self;

2. the Disorder Identity involves a shared understanding that typical group members are currently experiencing “symptoms”; and
3. (in the case of homeodynamic mechanisms), the social-comparative context changes in a way that systematically alters the group’s normative beliefs about “who we are” and “what we feel/think/do”, in contrast to a salient outgroup.

Disorder Identities as a means to adaptively modify the social environment

The arguments presented so far begin to explain how SCT can provide a social psychological foundation for externalist HPC similarity-generating mechanisms and homeodynamic mechanisms. But it is necessary to show not just in principle how these social-psychological mechanisms work, but to elaborate the sort of shared beliefs and theories about disorder that can drive stability and change in symptoms. The central argument that is presented here is that lay psychiatric theorising will explain much of the variation in these identity-based dynamics. This is an extension of Haslam’s (2005) folk psychiatry framework, reviewed in the previous chapter, according to which various dimensions of folk psychiatric beliefs have a distinct social-cognitive underpinning. Briefly again, folk psychiatric beliefs can involve normalising, medicalising, moralising, or psychologising the putatively psychiatric behaviour. Normalising relates to the judgement that a set of behaviours reflects normal human variation, medicalising the claim that a behaviour is attributable to internal causal processes beyond intentional control, moralising the view that a behaviour is intentional (and thus worthy of praise or blame), and psychologising the view that a behaviour is only partly under a person’s intentional control.

The claim developed presently is that rather than just viewing these folk psychiatric dimensions, and their underlying attributional logic, as outcomes of individual cognitive modes, these dimensions can be understood as fundamental drivers of inter-group dynamics. As Haslam notes, folk theorising about causation and intentionality saturates lay discussion

of mental disorder. This is not surprising given that these judgements govern how the behaviour of people with mental disorders is understood, the perceived blameworthiness of their behaviour, and people's perceived responsibility and motivation to change. Given what is at stake, it is little wonder that disorder-identity based groups might have strong views about these dimensions and seek to challenge competing folk psychiatric beliefs. If successful, as an act of minority influence, group members can carve out a viable social-ecological niche for explaining and justifying why group members behave and feel as they do, and how they should be treated.

In line with this, SIT predicts that the perceived legitimacy of intergroup status relations is a central determinant of whether or not individuals will accept or challenge the status quo. SIT suggests that a shared perception that intergroup status relations are illegitimate will tend to elicit responses at the intergroup end of the interpersonal–intergroup continuum. At the extreme, this can even manifest as violent social conflict—as in the example discussed in the previous chapter, when fibromyalgia lobby groups sent death threats to journal editors who published findings demonstrating the efficacy of a psychological intervention for their disorder, because of the two-step inference that the findings suggest that the disorder is a) not biological and therefore is b) blameworthy.

But more commonly, these reactions against dominant opinion can be understood as more nuanced minority influence attempts. As discussed in the previous chapter, certain anti-stigma campaigns attempt to challenge outgroup stereotypes that conditions like depression are a sign of moral weakness, by actively arguing that “depression is a disease like any other”. In contrast, the neurodiversity movement seeks to challenge societal and institutional beliefs that conditions like autism are diseases, rather than normal human variation to be celebrated. Some other groups claim that putatively disordered behaviour is actually under intentional control and even a sign of moral virtue. When pro-anorexia groups promote

mantras like “remeasure, reweigh, try harder”, these groups are not merely providing “thinspiration” for group members, they are articulating a distinctive theoretical understanding to the world that anorexia is a *choice*, with the implication that being thin demonstrates ingroup members’ superior moral virtue. A less widely controversial example of the same kind of strategy includes proponents of “death with dignity”. Again, the argument is that actively wanting to die is not a sign of psychopathology, or immorality, but rather a legitimate, and even honourable, moral decision. These examples are important because they challenge the notion that group-based self-perception necessarily entails an essentialist cognitive mode.

Table 2 summarises how quite different norms, and associated lay psychiatric theories, can be thought of systematically in terms of whether the ingroup perceives the valence of the “symptoms” to be positive or negative, and whether the condition is malleable or fixed. Importantly, the table also shows how it is that certain outgroup beliefs may threaten the ingroup representation. It is predicted that the process of context-dependent self-stereotyping will help explain how it is that these outgroup representations polarise the ingroup and thereby cause group members to hold these normative understandings more strongly. Of course this may play out over varying time-scales and in varying contexts. Even in seemingly interpersonal interactions between therapist and client, much of what is referred to as “resistance” (Beutler, Harwood, Michelson, Song, & Holman, 2011) can be understood as a manifestation of competing, group based, lay psychiatric theories.

Morgellans syndrome provides an interesting example (see Koblenzer, 2006). Morgellans is a contested illness that emerged in 2002 after a woman claimed to notice a fibrous substance emanating from her son’s skin. She noted similarities between her son’s symptoms and a condition, first described in the 17th century but undocumented since, called Morgellans disease. The woman subsequently established a website called the Morgellans

Research Institute and since then over 14,000 families in 45 countries have registered on the website indicating they believe they have the condition. Although sufferers are adamant that the condition is caused by an infectious parasite, almost all dermatologists and psychiatrists know Morgellans as Delusional Parasitosis (DP). Despite this, the group has been remarkably successful in carving out a viable niche for their community. Sufferers lobbied the United States congress and secured a \$500,000 research grant from the Centre for Disease Control for further research into their condition. Despite this research subsequently demonstrating that the condition is psycho-genic, lobbyists rejected these findings and continue to promote fringe academic journals and doctors that promote the condition. Sufferers argue that “like those who suffer from breast cancer and AIDS, we merely want appropriate resources devoted to our illness” (Chertoff, 2005, p. 60).

Morgellans identifiers continue to feel outraged that the mainstream medical community rejects their biological account of their condition. Psychiatrists note that when patients feel dismissed by medical practitioners, they commonly react by scratching their skin more compulsively—the physical wounds from which provide further “evidence”, to themselves and others, that their condition is physical. In other words, being a “Morgie” (as identifiers call themselves), in the face of outgroup threats questioning the biomedical legitimacy of the disorder, increases DP symptomatology and therefore seems to count as a positively homeodynamic mechanism. Other contested disorders, such as fibromyalgia, Dissociative Identity Disorder and chronic Lyme disease (Sigal & Hassett, 2002), display a very similar pattern of patient-clinician dynamics, a phenomenon dubbed recently as “iatrogenic doubt” (Floris & McPherson, 2015).

Of course the more ambitious question is whether these situational intergroup dynamics can really explain the sort of long term variation in a wider range of psychiatric syndromes, as implied by Hacking’s (1999) “looping effects”. As Hacking (1999) notes,

direct empirical evidence is limited in large part because of the challenges in empirically operationalising the looping effects thesis. However, theoretically it is quite plausible from a social identity perspective that chronic outgroup threats may increase symptomatology in a sustained way. For instance, various mental health anti-stigma campaigns contend that conditions like depression are “a disease like any other” and have been shown to reduce certain kinds of received stigma, but only at the cost of increasing sufferers’ prognostic pessimism, reducing their sense of self-efficacy, and capacity to manage their condition (Kemp, Lickel, & Deacon, 2013). To the extent this kind of prognostic pessimism exacerbates symptoms, as has been shown to be the case for depression (Abramson, Metalsky, & Alloy, 1989; Alloy, 1988) and anxiety (Dozois & Westra, 2005), then this too would seem to count as a positively homeodynamic, identity-based mechanism, and a clear example of a “looping effect”.

Table 2: *Disorder Identity Adaption Strategies to modify the social environment*

Ingroup beliefs about valence of signs / “symptoms”	Ingroup beliefs about malleability of signs / “symptoms”	Group aim(s)	Prototypical group norms	Ingroup lay psychiatric theory (Haslam, 2005)	Examples	Perceived threats from outgroup (“normals” and “experts”) predicted to strengthen ingroup beliefs
negative	malleable	seek outgroup support to change	"we need support and acceptance to grow and change in our own time"	psychologise symptoms	grieving process	medicalising, negatively moralising
negative	fixed	seek bio-medical support from outgroup to cure or palliate.	"we need support to live well in spite of our bio-medical condition" "we need biological research and biological cures"	medicalise symptoms	anti-stigma campaigns: "our condition is a disease like any other"	normalising, psychologising, negatively moralising
positive	malleable	seek outgroup <i>moral</i> acceptance of behaviour	"our behaviour is a legitimate moral choice" "our behaviour is virtuous"	normalise / positively moralise behaviours	death with dignity movement	pathologising, negatively moralising
positive	fixed	seek acceptance that signs and behaviours are normal human variation	"our behaviour cannot change and should be a celebrated as part of human diversity" "‘treating’ our condition is unethical"	normalise / biologise behaviour	gay rights activism (c. 1960 US) neurodiversity movement.	pathologising, negatively moralising, psychologising

Using Disorder Identity Adaption Strategies to modify the self

The adaption strategies discussed so far have focused on how it is that Disorder Identities can be a means for collectively contesting dominant lay psychiatric understandings and thereby modifying the social environment, and how in turn this contestation might actually amplify symptomatology as a function of the comparative context. In terms of the individual to inter-group continuum within SIT, these responses are towards the social change / inter-group end of the continuum. These responses depend on a strong Disorder Identity, having a shared perception that status relations are illegitimate and unstable, and using minority influence to attempt to, among other things, change outgroup lay psychiatric theories.

But this leaves open a range of complementary adaption strategies that do not necessarily involve (merely) changing the social environment, but rather involve using Disorder Identities to adaptively modify and control the self. Social creativity strategies are perhaps the most well documented strategy of this sort (Lemaine, 1974; Tajfel & Turner, 1979; Platow et al., 2014). To reiterate, these strategies are predicted to occur to the extent that leaving a low status group as an individual (i.e., individual upward mobility) is perceived to be unviable, and status relations are perceived to be relatively stable or legitimate. Under these circumstances it is predicted that group members may, for example, find alternative dimensions of comparison on which to base social comparisons (e.g. sad but wise), or by positively redefining the valence of a group-defining dimension (e.g. autistic and proud).

However, these identity management strategies also need to be understood as dynamic processes—and so must play out *over time*. Perceiving a social category does not just involve the perception that a category describes a static snapshot of the properties of a social group. Social category representations are invariably replete with narratives about not just

who we are, but who we have been (Liu & László, 2007) and who we will become (Bain et al, 2013).

This temporal dimension to social identities has received much less theoretical attention within the social identity tradition (Brown & Middendorf, 1996; Redersdorff & Guimond, 2006), but in many ways is implicit within SCT. According to SCT, the function of self-categorization is not just to define who we are but to make meaningful predictions (i.e. of the future) and explanations (i.e. of our past) (McGarty, 1999). For reasons mentioned earlier in this chapter, merely relying on manifestly observable similarities and differences between stimuli at a point in time is inadequate to constrain this categorization process. In part, this is because patterns of covariation do not make it clear which similarities and differences to attend to, but also because manifest similarities and differences are often deeply misleading. Category representations are therefore full of less obvious dispositional properties that (probabilistically) explain how category members will behave in certain counterfactual scenarios.

To make the point more concrete, the fact that I might wake up one day and no longer feel symptoms of depression is no guarantee that I will suddenly believe I am no longer a depressive. Why? Because I have a theoretical understanding about the nature of depression which means I understand this is just not how depression “works”. In the language of SCT, such a situation may display adequate “comparative fit” for the proposition “I am not depressive”. But this proposition displays poor normative fit (given my theoretical knowledge that depression seldom disappears overnight, and that I was feeling chronically depressed only recently) and my *readiness* to perceive this situation in these terms (given, for example, my longer term identification with the category “depressive”).

But this example also highlights the exquisite theoretical tension between comparative and normative fit and the intimate connections between categorization, attribution and social

influence. As touched on in the previous chapter, classic accounts of attribution (Kelley, 1967) show how people's causal inferences stem from the patterns of covariation between behaviour, person and situation. By contrast, evidence reviewed by Malle (1999) shows that attributional inferences are much more strongly influenced by information about people's intentional states than by patterns of covariation. However, from a self-categorization perspective, it is the interaction between covariation information (comparative fit) and theory (normative fit) that governs people's attributional inferences (see also Ahn, Kalish, Medin, & Gelman, 1995).

Shampoo advertisers know this well. Adverts for at least one premium product contend that by using their shampoo, shiny hair “won’t happen overnight, but it will happen”. It is just as well, because consumers who subsequently notice few if any changes in the shininess of their hair might otherwise wonder if they had wasted their money. Indeed there are a range of competing theories that might explain my dull, limp hair on day two: that the shampoo does not work (externalising), or the dispositional inference that my hair just does not work the way that supermodel hair works. The point is that compelling theories must not only explain the past, and predict the future, they must also explain away apparent disconfirmations. Specifically, lay psychiatric theorising needs to predict stability and change, but also account for instances where the normative fit is seemingly poor.

Merely holding a particular lay psychiatric theory as an isolated individual—my unique kind of depression works thus and so—profoundly limits the opportunities for lay theory generation and testing. One does not need to be a whiz in predicate logic to see that if the experience of others is to have any self-relevant implications, then the self must at least be part of the universe of discourse in which theoretical claims about category members are formulated. This is not to imply that individuals seek out groups because they want certainty about their condition (Hogg, 2000). On the contrary, groups often generate subjective

uncertainty—as when my symptomatic experience of a condition departs radically from the normative trajectory of prototypical group members. Again, this tension between the normative and comparative “fit” of my predicament can be resolved either by reinterpreting my symptoms so they align with the norm (“Maybe I’m actually feeling better (or worse) than I thought”—i.e. social influence), by explaining away the discrepancy (“I am just having a bad week, these setbacks are normal in depression”), or by recategorizing the self (“maybe I’ve got something else wrong with me”) or others (“maybe *they* are not really depressed”). The social identity analysis suggests that which of these resolutions occurs depends on a person’s long term identification with the category, the availability of cognitive alternatives (“what else could explain who I am?”), and the theoretical understandings of a condition that are available.

This point has significant implications for how it is that shared disorder representations underpin the dynamics of stability and transition into and out of disorder. Disorder identities need to provide a normative logic, informed by lay psychiatric theories, for understanding how a particular disorder “works”: one that explains and predicts stability and change in symptoms over time. Of course these narratives cannot be wildly at odds with the lived experience of group members and will be constrained by whatever other “brute” causal mechanisms underpin an HPC disorder. A group leader who constructs a lay narrative that is at odds with group members’ experience, and the brute constraints imposed by a disorder, is vulnerable to being challenged. In these circumstances, other ingroup members are likely to be perceived as better representatives of “us” to the extent they can articulate a narrative that “fits” with our lived experience and creates a positive vision for our future (Haslam, Reicher, & Platow, 2010).

The challenge that has been set is to systematically document important dimensions of these different narratives in the context of various mental disorders and to map the possible

dynamic consequences of these narratives within an HPC framework. In Chapter 2, as well as homeostatic and homeodynamic mechanisms, four further kinds of HPC processes were identified: positively and negatively heterostatic processes; and positively and negatively heterodynamic processes. I do not attempt to explain these lay psychiatric narratives comprehensively. Rather, the focus will be on the two lay dimensions of valence and malleability, given the crucial role they play in motivation and behaviour as discussed in Chapter 3, and to claim that these two dimensions capture important sources of variation in how lay narratives are constructed and how these in turn might drive HPC dynamics.

Disorder Identity Adaption Strategies as heterostatic mechanisms

Positive and negative heterostatic processes are those processes that inhibit transition into and out of an HPC category respectively. In Table 3, I document how Disorder Identity groups can promote very different lay theories about the valence and malleability of a condition and that these can, in turn, have either positively or negatively heterostatic consequences that tend to stabilise groups. The valence/malleability mapping also clarifies why it is important to consider those disorder-related identities in which group members may not currently meet criteria for an HPC disorder, but nevertheless affect HPC dynamics. As mentioned earlier, most members of Alcoholics Anonymous and other, similar, 12-step programs, would not meet criteria for substance-related disorders. Nor is there currently any empirical evidence to suggest that current and former substance abusers share some underlying disease essence (Hall, Carter, & Forlini, 2015). Despite this, AA programs provide a normative understanding that members *are incurably diseased* and that committing to the group is the only way to prevent group members from reverting to type. The Social Identity Approach suggests that this requires group members not just to identify as alcoholic, but that group members need to accept the group's normative logic about the theoretical nature of alcoholism (its essentialist underpinnings) and the role of the group in maintaining

sobriety. Leaving aside whether this normative logic is “true”, there is empirical evidence that being an AA member can prevent relapse (Ferri et al., 1996). This suggests, then, that social identification with this Disorder Identity, and so with its associated logic, has positively heterostatic consequences. While to some it might seem counter-intuitive that a wrong-headed lay theory can “work”, the point is simply that these folk theories can have powerful motivational and behavioural consequences and so alter the dynamics of HPC disorders.

However, conceivably, it may not be necessary for Disorder Identity groups to adopt fixed and negative beliefs in order to prevent transition into HPC defined disorder. A fascinating example of this, in the case of substance abuse disorder, has been the development of so called “Moderation Management” groups (Lembke & Humphreys, 2012). These are mutual self-help groups that reject the AA notion that conditions like alcoholism are diseases and instead claim that problem drinking is a learned behaviour and that group members can support each other to drink safely and in moderation. Binges can therefore be attributed not to an enduring disease essence, but rather are explained away as a mere “lapse”. Perhaps not surprisingly, given their very different lay theories about drinking behaviour, AA and Moderation Management groups are fiercely antagonistic towards each other (Hodgins, 2005). But despite this, there is evidence that abstinence and moderation management can both be effective in reducing relapse into problem drinking (Hester, Delaney & Campbell, 2011). The implication of these findings for the Disorder Identity Analysis presented here is that what may matter most in terms of symptom management is not so much the truth or falsity of the lay theory, but whether group members perceive that the lay theory provides a compelling narrative for explaining group members’ symptom experience, whether group members psychologically identify with the group, and whether the groups’ norms drive relevant behaviours and motivations.

The two examples discussed so far are of positively heterostatic mechanisms that prevent relapse into the disordered state. However, it is also predicted that some Disorder Identity groups can have *negatively* heterostatic consequences—and make it hard to leave the HPC category. The two possibilities documented in Table 3 include social creativity strategies and what might be called moral discipline strategies. Social creativity strategies, that positively redefine “symptoms” or define the group on alternate dimensions of comparison, may end up serving as a barrier to leaving the HPC category. Suggestive evidence for this process is provided by the experience of people within the deaf and signing community who choose not to receive cochlear implants, either for themselves or their children (Most, Wiesel, & Blitzer, 2007), or the previously cited example of individuals with achondroplasia (dwarfism) in the United States (Branscombe et al., 2012). Importantly, these positively defined Disorder Identities evolved in a context where, at least initially, there were no technologies available to overcome a condition, and so were in that sense “fixed”. But tellingly, when technological advance presented the opportunity to leave the group, the culture that had been established around the category became a barrier to leaving. To reiterate, this is not to insinuate that this consequence is “bad”, but merely to highlight how identities are historically situated and that technological change continually generates opportunities to forge new understandings of a Disorder Identity as well as creating threats to existing ones.

In contrast to this class of social creativity-based stabilising strategies, moral discipline strategies prevent ingroup members from leaving by exerting moral influence. The normative logic here is very different. There is a shared understanding that a behaviour is, and always was, an intentional behaviour and does not reflect any deep, biologically-inherent properties of category members. In these circumstances, those who threaten to leave the group are chastised as failures, and shamed for lacking self-discipline. Again, the example of

moral shaming, so called “fat shaming”, within certain pro-anorexia groups seems apposite here (Lupton, 2013).

There are two important points of clarification about these heterostatic Disorder Identity norms. First, there may be considerable variation in the extent to which these norms are formally articulated, the level of consensus about the norms, and their level of endorsement. Groups which attempt to influence an outgroup to agree that a behaviour can be a positive moral choice need not adopt norms that have heterostatic consequences. “Death with dignity” groups do not argue that people who choose not to take their own life are cowardly, just that those who do make this decision should not be understood as being mentally ill or dishonourable. The second point of clarification is that flexible combinations of these normative understandings are possible, and indeed theoretically implied by the flexibility of the self-categorization process. As discussed in the previous chapter, some fibromyalgia group members privately note that in some ways they have come to positively value being part of the fibromyalgia community and would not seek to return to their previous way of life (social creativity, negatively heterostatic consequences). This seems to suggest a positive, fixed understanding—but at the same time, fibromyalgia groups commonly promote “adaptive pacing” to help disorder sufferers manage their condition (palliative strategy, (possibly) positively heterostatic consequences). This implies that the norms within this disorder group are something of a hybrid between the social creativity and palliative strategies shown in Table 3.

Table 3: *Four kinds of heterostatic (stabilising) norms.*

Valence of “disorder” (from perspective of ingroup)	
Fixed	Positive
	Strategy: social creativity
	Normative logic: “who we are can’t change but we define ourselves positively in other ways / positively celebrate our condition”
	Likely HPC consequence: negatively heterostatic
	Dominant lay psychiatric theories: essentialise and depathologise condition.
Malleable	Negative
	Strategy: symptom reattribution
	Normative logic: “our future is not defined by our history of disorder, even if we ‘lapse’ or engage in “symptomatic” behaviour in moderation”. “lapse is not the same as ‘relapse’”
	Likely HPC consequence: positively heterostatic
	Dominant lay psychiatric theories: psychologising, normalising
Malleability of “disorder” (from perspective of ingroup)	
Fixed	Strategy: moral discipline / shaming
	Normative logic: “we proudly choose our behaviour and those who leave our group lack self-discipline”
	Likely HPC consequence: negatively heterostatic
	Dominant lay psychiatric theories: positively moralise condition.
	Example: Pro-anorexia Groups
Malleable	Strategy: symptom reattribution
	Normative logic: “our future is not defined by our history of disorder, even if we ‘lapse’ or engage in “symptomatic” behaviour in moderation”. “lapse is not the same as ‘relapse’”
	Likely HPC consequence: positively heterostatic
	Dominant lay psychiatric theories: psychologising, normalising
	Examples: “Moderation Management” groups

Disorder Identity Adaption Strategies as heterodynamic mechanisms

So far, I have shown how SCT predicts that the process of depersonalisation will drive homeostatic mechanisms, context-dependent self-stereotyping will drive homeodynamic mechanisms, and differing normative theories will drive heterostatic mechanisms. The remaining piece of the conceptual puzzle is to show how norms might drive heterodynamic mechanisms. This relates to the question of how it is that group processes facilitate transition into and out of psychological groups, and how this in turn could influence flows in and out of an HPC defined category. Such identities have previously been described as “transitional identities” (Kellog, 1993), in the sense that there is an expectation that, over time, group members transform themselves towards some end state, informed by a theoretical narrative. These narratives provide a lay theoretical account explaining how it is that group members can become a different person (at either a surface or “deep” level) in the future compared to who they were in the past (Howard, 2008).

The social identity literature examines in some detail how it is that group leaders can facilitate these kinds of identity transition. For example, leaders can act as “entrepreneurs of identity” by appealing to, and constructing, a shared identity that mobilises collective support for a cause (Reicher, Hopkins, Levine, & Rath, 2005). This often involves constructing a lay theoretical framework that articulates who “we are” and in a way that draws attention to which surface-level similarities and differences are important. In a clinical context, the basic psychological processes that facilitate transition into a Disorder Identity group are not fundamentally different. Again, what is needed is a theoretical narrative that explains who group members are, who they were, and who they could become. A possible conceptual mapping of these strategies is outlined in Table 4, again showing how these could play out in different ways, as a function of perceived malleability and valence.

Table 4: *Four kinds of heterodynamic (transitional) norms.*

Valence of “disorder” (from perspective of ingroup)	
Fixed	Positive
	Strategy: coming out proud
	Normative logic: “the group is here to support you to come to terms with who you essentially are. People feel a sense of relief that they no longer have to repress their true self, and can even come to have a sense of pride in who they are”
	Possible HPC consequence: negatively heterodynamic
	Dominant lay psychiatric theories: essentialise and depathologise condition.
Malleable	Example: Coming out as “fat”(Saguy & Ward, 2011)
	Strategy: aspirationalism
	Normative logic: “if you work hard, you too could become like us”
	Possible HPC consequence: negatively heterodynamic
	Dominant lay psychiatric theories: positively moralise condition.
	Example: Pro-anorexia groups and “thinpiration”
	Strategy: collective mobility
	Normative logic: “we can grow and change and thereby overcome our condition”
	Possible HPC consequence: positively heterodynamic
	Dominant lay psychiatric theories: psychologise
	Example: Cognitive therapies for depression
	Strategy: accept and control
	Normative logic: “accepting that we are sick is the first step towards managing the symptoms of our disease”
	Possible HPC consequence: positively heterodynamic
	Dominant lay psychiatric theories: pathologise
	Example: Alcoholics Anonymous

The notion that certain social groups could create the conditions that cause individuals to transition from a non-disordered state to a “disordered” state is controversial and difficult to demonstrate empirically, but is a clear implication of Hacking’s (1999) “looping kinds” argument. For example, Hacking notes that the number of people with Dissociative Identity Disorder increased from 200 to 20,000 within a decade and he partly attributes this increase to this set of behaviours increasingly becoming a distinctive, and socially understood, “way of being”.

However, most groups would reject the notion that they foster such transitional identities. Obvious exceptions to this are certain pro-anorexia and pro-bulimia groups which construct negatively heterodynamic “thinspirational” identities to the extent they can articulate the promise that, through effort and self-discipline, it is possible for outsiders to become part of the group, with all of the alleged positive status that comes with this. Again, this strategy reflects an underlying premise that this behaviour is both positive and malleable.

Examples of negatively heterodynamic transitional identities, informed by views that the focal behaviour is *fixed* and positively valenced, are more difficult to identify. This is in part because such groups are likely to disavow the proposition that identity processes play a causal role in changing behaviour. Instead, such groups are likely to claim that identifying with a condition is just the process of coming to terms with a pre-existing, underlying condition. Perhaps the clearest examples are modelled on the process of “coming out proud” within queer communities. While homosexuality (since 1973) and transgenderism (since 2013) are no longer considered disorders, there is evidence from representative population surveys that the prevalence of transgendered behaviour (Zucker, & Lawrence, 2009) as well as homosexual behaviours, feelings and identification (de Visser et al., 2014, Johnson et al., 2013) have increased over time, and in ways that are likely to reflect changes in orientation and behaviours rather than an increasing willingness to disclose (Prah et al., 2013).

Nevertheless, from the perspective of those who hold an essentialised view of these disorders, this is likely to be understood not as an increase in the prevalence of homosexuality or transgenderism, but rather an increase in the number of people who accept their “true, underlying” feelings.

But from a scientific perspective, sexual orientation and gender identity is multifactorial: having both a biogenetic basis (Bailey et al., 2016) as well as being shaped by environmental, including social, influences (Diamond & Rosky, 2016). Diamond (2012), for example, summarises extensive longitudinal evidence for fluidity in sexual orientation over time, as well as evidence that women brought up by lesbian parents are more likely than matched controls to experience same-sex attraction (Gartrell, Bos, & Goldberg, 2011). Diamond concludes that the most likely explanation for this is that by growing up in a positive social climate of acceptance, individuals “may have been more willing and able to consider—and to positively evaluate—their own propensity for same-sex sexuality (p. 371). This is clearly consistent with Hacking’s (1999) looping kind thesis. The further particular claim being made here is that in order to create this “positive climate of acceptance” an identity-based theoretical narrative is required, such as is implied by “coming out proud”, in order to facilitate this transition.

There is evidence that “coming out proud” is a strategy that has been adopted in relation to various conditions currently classified as disorders, including binge eating disorder and “fat pride” (Saguy & Ward, 2011), schizophrenia (Corrigan, Kosyluk, & Rüsch, 2013) and autism spectrum disorder (Davidson & Henderson, 2007). It should be noted that there are important differences between these conditions that are likely to create differences in emphasis in how a “coming out” strategy is applied (Corrigan et al., 2009). Sarrett’s (2016) qualitative research provides insight into the normative logic underpinning the “coming out” approach in the context of an online autism forum. Within the online community that Sarrett

researched there was a shared view that autism should be neither normalised nor pathologised. However, forum members had polarised views about the process through which individuals could legitimately *claim* to be autistic. Some forum members insisted that a professional diagnosis was required before participating actively in the forum, but others felt individuals who self-identified as autistic were welcome, and noted that the “coming out process” takes time, may be characterised by initial feelings of ambivalence and uncertainty and therefore symptoms may not be properly recognised by clinicians. This qualitative evidence is consistent with the broad argument that even groups with a fixed understanding about the nature of a condition can be proactive in supporting people to transition into the group.

In contrast to these potential examples of negatively heterodynamic Disorder Identities, there are a number of examples of transitional identities that seem to support transition from active clinical disorder to a non-disordered state (i.e. are *positively* heterodynamic). As mentioned on a number of occasions already, AA programmes encourage (potential) group members (particularly in the early stages) to accept that they are alcoholic (“admit we are powerless”) and that failing to admit this is denial. In this instance, the ingroup understanding is that conditions such as drug and alcohol addiction are negatively valenced and fixed. On this view, what changes is group members’ acceptance of their “true nature”, and with this comes the capacity to control their symptoms. As mentioned in earlier chapters, ironically this has the consequence that individuals can transition from a clinically diagnosable alcoholic state to a non-clinical state (from the third person, professional perspective) by psychologically identifying as an alcoholic (from the first person perspective).

Positively heterodynamic transitional identities, involving negative and fixed beliefs about a condition, may also be common in other therapeutic contexts. The lay theoretical

logic of certain pharmacological treatments is one such example. When selective serotonin reuptake inhibitors first became one of the front-line treatments for Major Depression, it was common for pharmaceutical companies to provide elaborate, but consumer-friendly, educational material about how these products “worked”—complete with pictures of synaptic clefts awash with happiness-inducing serotonin molecules correcting a prior biological imbalance that purportedly causes depression (Lacasse & Leo, 2007). This kind of marketing promotes a fixed theory of depression (or at least “unintentional” in the sense outlined by Malle, 1999), since it implies that the condition would persist in the absence of this external bio-chemical intervention (even though this so-called “bioamine hypothesis” has been repeatedly, empirically debunked, see Valenstein, 1998, Ionnidis, 2008). However, if at least 80 per cent of the treatment effect of anti-depressant medication is a placebo response (Khan & Brown, 2015, Kirsch et al., 2002) then it makes good sense why pharmaceutical companies would seek to actively promote this as a lay theoretical understanding of depression. Doing so perpetuates a powerful, and therapeutically consequential, social expectation about how these drugs work (Barrett et al., 2006). Consistent with this, Walsh, Seidman, Sysko, and Gould (2002) found that the placebo response to anti-depressants has increased by approximately 7 per cent each decade since 1981 (see also Kasper & Dold, 2015). Walsh et al. note that this trend can be observed even after controlling for a range of other explanatory variables, such as variation in the characteristics of baseline samples and study protocols. The authors concluded that the effect is most likely due to changing social expectations about the nature of depression and its causes and treatment. Further consistent with this, there is evidence that lay beliefs about the cause of depression are becoming increasingly biologically-based over time (Blumner & Marcus, 2009), and such beliefs have been shown to increase the perceived desirability and efficacy of antidepressant medication (Kemp, Lickel, & Deacon, 2014). To be clear, these findings are merely consistent with, but do not

directly demonstrate, the further claim made here that this is a Disorder Identity-based process. However, these findings are certainly very consistent with the claim that, for those taking anti-depressants at least, a depressive identity, with a negatively valenced and biogenetically-caused normative understanding could serve as a positively heterodynamic mechanism—even if, as many have noted (Kvaale et al., 2013, Schultz, 2015), such a theoretical account has various, more problematic consequences at a societal level.

So far, I have considered positively heterodynamic Disorder Identities involving negatively valenced and fixed beliefs about the nature of a disorder. A relatively under-explored possibility, however, is that certain Disorder Identity related groups can promote the idea that a disorder is negatively valenced, but malleable. This is most likely to involve Disorder Identity group members invoking a lay psychiatric theory of psychologising (Haslam, 2005) to explain the nature of their condition. This is because such a theory implies that Disorder Identity group members have only partial insight into the causes of their behaviour (and so helps mitigate the inference that group members are blameworthy). However, such a theory also suggests that group members are likely to believe that psycho-social interventions could ultimately be effective in overcoming their condition (Ahn, Proctor, & Flanagan, 2009, Lebowitz, 2014). Indeed this is the logic that applies to many therapy programs in the cognitive and behavioural tradition which aim to promote a malleable theory of emotional distress (Kneeland, Dovidio, Joormann, & Clark, 2016). In a theoretically similar vein, various popular books (Doidge, 2008) promote the notion that the brain is plastic and so “real”, enduring, change is possible—in much the same way as it is possible to learn a skill.

From the perspective of SIT, this Disorder Identity Adaption Strategy corresponds to the social identity management strategy of *upward mobility* (i.e. from a low status/negatively valenced group to a high status group). However, whereas this strategy is typically

understood as an individualistic response to being in a low status group with a permeable intergroup boundary, it is claimed here that this is in important respects a collective process. In the next chapter, I will elaborate this claim in more detail and argue that this strategy of collective upward mobility provides a useful starting point for empirically testing the broader Disorder Identity Analysis.

Summary: Disorder Identity Adaption Strategies as a dynamic system

In the foregoing analysis I have at last mapped out an array of possible Disorder Identity Adaption Strategies, summarised in Table 5 below, and begun to systematically describe some of the extraordinarily diverse ways that people with disorder can “recover”. Throughout this analysis I have endeavoured to avoid normative judgements about these strategies. Rather, I have attempted to describe what Disorder Identities might mean from the perspective of people who inhabit them and how shared understandings of disorder can form a coherent set of inter-related beliefs about the causal basis of disorder and the features of the social context in which they are experienced. Of course the framework does not purport to be exhaustive and nor does it take into account how lay understandings of mental disorder may be organised in very different ways in different cultures. A notable lacuna, for instance, is the neglect of spiritual understandings of disorder, a point also noted by Haslam (2005) in his discussion of lay psychiatric theorising.

However, the further claim is that these Disorder Identity Adaption Strategies are not merely of descriptive, sociological interest, but that the social-psychological processes outlined in the Social Identity Approach could be used to make predictions about how beliefs and behaviours can become socially shared ways of understanding the self. Specifically, SCT provides a falsifiable account of how socially shared self-category representations get in under the skin and become psychologically salient in context. Importantly, SCT also provides a well-established theoretical basis for understanding how beliefs and behaviours

can dynamically change. This includes depersonalised self-perception leading to ingroup homogeneity and context-dependent self-stereotyping leading to polarisation of beliefs and behaviours away from contextually relevant outgroups. In turn, these psychological processes correspond to homeostatic mechanisms and homeodynamic mechanisms respectively (as shown in Table 6).

The broader Social Identity Approach also provides a general account of how groups of people strive to construct (or aspire towards) a positive social identity, while taking into account the structural features and constraints of both their environment and their embodiment. It is in this sense that all of these strategies can be understood as fundamentally recovery-oriented and adaptive from the perspective of group members—even if some of the strategies are more widely viewed as dangerous, or socially problematic, or that similar ends could be better achieved through different means. From the perspective of group members, sometimes this adaptive process can be best achieved by seeking to change the broader social environment—for example, by seeking to influence outgroup members to accommodate, or change, or perhaps even depathologise “symptoms”. At other times, this will involve seeking to modify or maintain the self—including both the individual self and other (potential) ingroup members. As outlined at length, this involves group members constructing a normative logic and lay psychiatric theory that explains who we are, who we were and who we could become. In turn, it is argued that these narratives provide the basis for a variety of social identity management strategies that are predicted to drive heterostatic and heterodynamic mechanisms, as shown in Table 6.

It is important to emphasise again that the HPC and the Disorder Identity Frameworks that make up the Disorder Identity Analysis of Looping Kinds are not simply restatements of each other. To reiterate, people can perceive themselves to be disordered, but not meet clinical criteria for disorder; and conversely a person can reject the notion that their behaviour

is disordered but meet formal criteria for disorder. Similarly, people's lay psychiatric beliefs may or may not be consistent with empirical research evidence, but these beliefs may nevertheless have important behavioural, cognitive and emotional consequences that substantively alter clinical outcomes.

A further point to note is that the various strategies I have outlined should not be understood mechanically, nor as neatly and categorically distinct. What should be clear is that any particular Disorder Identity group may be constantly responding to changes both within the group and the social context. Moreover, there may be competing Disorder Identity groups associated with particular conditions, proffering very different understandings of how conditions come to be, as well as tensions and trade-offs within groups. At a macroscopic, epidemiological level, then, the morphology of a particular HPC disorder category might display apparent stability and relatively incremental net change in the stock and flow of category members, and in the nature and severity of symptoms that are displayed. But underlying this, the framework implies there is likely to be a thrum of meso-level activity, as different Disorder Identity groups, and subgroups, use these strategies to compete, react and reshape themselves, each other, and their wider social context.

It is important to note that the emphasis here has been on how disorder identifiers change their behaviour as a consequence of having a shared awareness of a disorder category. These are changes that are theorised to occur while holding constant the thresholds for clinical diagnostic categories, and the diagnostic cultures and decisions of clinicians. However, as Haslam (2016) (*pace* Hacking, 1999) suggests, the scope of psychological and psychiatric concepts can be subject to "creep" such that, holding constant any changes in the behaviour of individuals, the boundaries of what counts as disorder can change. A more complete, and necessarily complex, account of the Disorder Identity Analysis described here, therefore, needs to take into account the identity dynamics that operate *within* professional

and academic communities. These clinical judgements about what is and what is not within the scope of a disorder category are likely to change. Sometimes these changes will reflect changes in technology and society, many of which are in turn caused by social activism on the part of Disorder Identity groups. While the processes underlying this “conceptual creep” are not discussed in depth here, the Social Identity analysis of social influence would likely provide a robust theoretical starting point for such an enquiry.

I also acknowledge the risk that in exploring the identity-based processes that could theoretically drive these dynamic changes in disorder, I may be interpreted as reinforcing stigmatising and disempowering attitudes towards people with mental disorder. For example, it may be perceived that these arguments serve to reinforce the view that people with mental disorder could choose to behave differently, and are therefore blameworthy, or that mental disorder is somehow not real because it can be causally influenced by social processes. This is certainly not my intent. As argued at length in Chapter 1, I presume that to varying degrees “brute”, internal causal mechanisms operate beyond conscious awareness and contribute causally (in ways not yet understood) to the experience of psychological distress in all its myriad forms. Informed by critical realism (Pilgrim, 2013), I further assume that these brute causal processes operate independently of human discourse and can be usefully studied empirically, albeit imperfectly. However, to reiterate, I do make the claim that the symptoms of disorder can also be causally influenced by an individual’s beliefs and behaviours and that these in turn are causally influenced and indeed constrained by the social-psychological groups to which individuals belong.

The implications of these points are important. Firstly, Disorder Identity groups are unlikely to develop lay psychiatric theories that are wildly at odds with whatever other biological-determined constraints a condition imposes. A Disorder Identity group informed by the shared experience of Autism Spectrum Disorder is unlikely to adopt a “collective

mobility” strategy, simply because the claim that Autism is a condition that can be readily “treated” and overcome is likely to be at odds with the lived experience of most group members. Secondly, individuals cannot straightforwardly choose which Disorder Identity groups they identify with (if any), nor choose as an individual to change the group norm. The Social Identity Approach does not imply an “anything goes” constructionism. Rather, an individual’s beliefs, behaviours, and feelings are constrained by the interaction between whatever brute mechanisms contribute to the disorder experience and the varying narratives provided by rival groups and subgroups that purport to explain the disorder. Arguably the main job of a good therapist is to deconstruct “unhelpful” narratives about the causes and consequences of distressing symptoms. That this therapy is fundamentally a social influence process does not imply that this does not cause real changes in people’s brains (Buhle et al., 2014), nor that those who do not experience symptom improvement are blameworthy.

A third point worth noting is that even if it was the case that groups or individuals could freely and straightforwardly “choose” to behave differently (which, to reiterate, I am not suggesting) it does not follow that they *should*. As many have persuasively argued, despite the rhetorical force of lay essentialism in countering stigmatising views, it does not logically follow that a person who, for example, declares that they intentionally choose to be gay, should automatically be chastised as immoral (Diamond & Rosky, 2016).

To anticipate a further potential criticism, I acknowledge that the proposed Disorder Identity Analysis is primarily informed by theoretical arguments. Direct empirical evidence to support the analysis is scarce. Most notably, very few researchers have measured Disorder Identity in terms of social identification (Adams et al., 1997, Jetten et al., 2011) or the group norms associated with Disorder Identities (Buckingham et al., 2013). Even the indirect evidence for the claims made within the Disorder Identity Analysis, while suggestive, are often scattered and reliant on qualitative and observational studies. An understandable

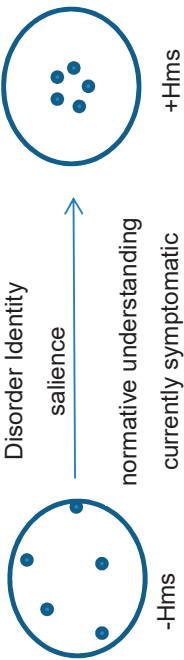
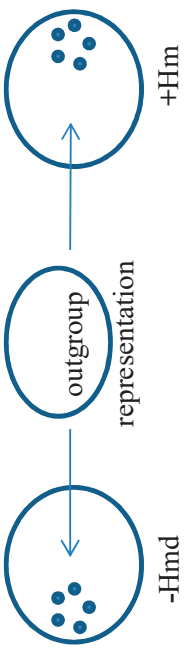
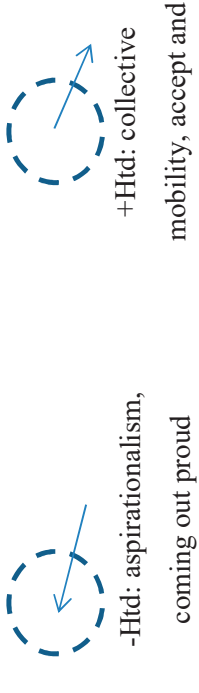
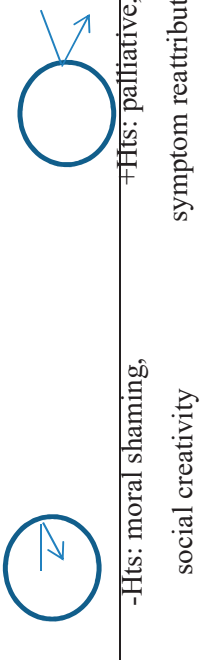
concern, then, is that proposing such a sweeping conceptual framework on the basis of such comparatively weak empirical evidence is, at best, premature.

This line of criticism needs to be considered in the context of a research domain that is widely claimed to be in crisis, by commentators from diverse perspectives both within and outside of psychiatry and clinical psychology (Frances, 2013, Kandel, 1998, Katschnig, 2010, Kindermann, 2014, Insel & Wang, 2010, Morgan, 2015, Pigani et al., 2015, Pilgrim, 2013). Calls to develop a “new paradigm” have become cliché. But interestingly, specific proposals for what such a new paradigm might look like are much harder to find (Bracken et al., 2012). In part, this is because any evidence for a new paradigm, at least initially, must be salvaged from old ones—even though this evidence is likely to be patchy and difficult to integrate. The more appropriate criteria for judging any new approach, then, should be whether the approach can generate testable hypotheses that are informed by theory. By this measure, it is hoped the Disorder Identity Analysis can be judged in a more promising light: if for no other reason than it would be useful to prove the predictions implied by the analysis, and Hacking’s broader “looping kinds” thesis, to be wrong.

Table 5: *Disorder Identity Framework: a taxonomy of Disorder Identity Adaption Strategies*

	Ingroup beliefs about valence of signs / “symptoms”	Ingroup beliefs about malleability of signs / “symptoms”	Adaption Strategy
Modifying the social world	negative	malleable	seek outgroup support to grow and change in own time
	negative	fixed	seek bio-medical support from outgroup to cure or palliate.
	positive	malleable	seek outgroup moral acceptance of behaviour
	positive	fixed	seek outgroup acceptance that signs and behaviours are natural human variation
Modifying the self	negative	malleable	symptom reattribution
	negative	fixed	palliation
	positive	malleable	moral shaming/discipline
	positive	fixed	social creativity
	negative	malleable	collective mobility
	negative	fixed	acceptance and control
	positive	malleable	aspirationalism
	positive	fixed	coming out proud

Table 6: *HPC Framework: Disorder Identities as HPC mechanisms*

HPC mechanism	Social identity process
1. Homeostatic mechanism 1.1. +Hms = positively homeostatic mechanism, increases the degree of similarity between HPC category members. 1.2. -Hms = negatively homeostatic mechanism, decreases the degree of similarity between HPC category members.	
2. Homeodynamic mechanisms 2.1. +Hmd = positively homeodynamic mechanism, increases the average severity of HPC symptoms. 2.2. -Hmd = negatively homeodynamic mechanism, decreases the average severity of HPC symptoms.	
3. Heterodynamic (transitional) mechanisms 3.1. +Htd = positively heterodynamic mechanism, facilitates outflow of HPC category members. 3.2. -Htd = negatively heterodynamic mechanism, facilitates the inflow of HPC category members.	
4. Heterostatic (stabilising) mechanisms 4.1. +Hts = positively heterostatic mechanism, prevents inflow of HPC category members. 4.2. -Hts = negatively heterostatic mechanism, prevents outflow of HPC category members.	

Chapter 5: Introduction to the Empirical Chapters

The task of testing all of the myriad empirical claims implicit within the Disorder Identity Analysis is substantial, and well beyond the scope of this thesis. However, in this chapter I narrow down on just one very specific Disorder Identity Adaption Strategy that I proposed in Chapter 4 relating to collective mobility. To reiterate, this strategy involves Disorder Identity group members developing a malleable, rather than a fixed, group-based self-stereotype that serves to motivate group-members to strive to overcome, rather than merely palliate symptoms. Depending on the nature of the condition, it is hypothesised that such a Disorder Identity Adaption Strategy would have positively heterodynamic and negatively homeodynamic consequences.

This proposal is perhaps one of the more controversial claims within the framework that I have developed. Whatever other putative benefits Disorder Identities might confer, it seems counterintuitive that it could be at all advisable for individuals to develop a Disorder Identity as a means to reduce symptomatology (i.e. be negatively homeodynamic) and perhaps even overcome that condition (i.e., be positively heterodynamic). Corrigan, Larson, and Ruesch (2009) highlight a commonly held view that such identities may well provide a basis for group members to give and receive support and counter received stigma, but there is also a risk that such groups inadvertently strengthen *self*-stigma, leading to what they call a “why try effect”. The point reflects a trade-off noted earlier: the commonly adopted group-based strategy of strategic essentialism has been shown to counter stigmatising views that the mentally ill are blameworthy, but only at the cost of inducing “prognostic pessimism” (Kvaale et al., 2013). Some researchers go a step further in suggesting that group-based perception is inevitably tied to holding more essentialist beliefs about the causes of a condition (Levy et al., 2001, c.f., Morton et al., 2009). I will therefore focus in on this issue

of collective mobility as outlined here, not only as a test of the broader viability of the Disorder Identity Analysis, but because this would seem to be a particularly critical test.

Even within SIT, it is not at all clear that upward mobility (from a low status to a high status group) could plausibly be construed as a collective rather than an individuating process. In SIT terms, perceived boundary permeability between groups (which can be equated in this context with the perceived malleability of a condition) is likely to reduce low status group members' identification with the low status group and increase the likelihood of low status group members adopting an *individual* upward-mobility strategy (Ellemers, van Knippenberg, De Vries, & Wilke, 1988; Tajfel & Turner, 1979; Tajfel, 1974). Conversely, "... the idea that 'passing' individually from one's own group to another is impossible or extremely difficult causes ... people from that group to feel and act in unison" (Tajfel, 1974, p. 106).

These theoretical predictions have been supported in experimental manipulations of group boundary permeability in laboratory created groups (Ellemers et al., 1988; Ellemers, Van Knippenberg, & Wilke, 1990; Jackson, Sullivan, Harnish, & Hodge, 1996). Ellemers et al. (1988), for example, found in an experimental laboratory study that when low status group members were provided feedback that boundaries between the low and high status group were relatively permeable, participants' identification with the low status group was lower than when the boundaries were presented as impermeable. There is also experimental evidence for the reverse causal sequence. Ellemers, Spears and Doosje (1997) manipulated levels of group identification with a low status group using a bogus pipeline procedure and found that high levels of identification caused low status group members to perceive the ingroup as more homogeneous and to be more reluctant to leave the group as an individual, even when presented with evidence that group boundaries were, in fact, relatively permeable.

Together, these findings suggest that identification as a member of a low status group can be both a cause and a consequence of perceived boundary impermeability and that upward mobility is an inherently individuating process. But this theoretical claim is precisely at odds with the collective mobility strategy outlined in the Disorder Identity Framework, according to which upward mobility from a low status group could in some sense be understood as a collective process.

This theoretical tension stems in part because boundary permeability within SIT is invariably considered categorically, as either permeable or impermeable, and epistemically as an objective reality constraint that is uncontroversially perceived or obviously given by the context (Ellemers et al., 1990). This is true for many bio-medical illnesses that are defined essentially, and for which the prognosis is clear, such as the common cold or terminal cancer. For mild, transient conditions, such as the common cold, individual recovery is almost guaranteed and so it would not be expected that group-processes play a significant psychological role in the process of symptom recovery. For more severe conditions, for which cure or permanent symptom remission is unlikely, it would be expected that group members adopt various identity management strategies that are predicted to arise when intergroup boundaries are perceived to be impermeable—such as positively defining the group on alternative dimensions of comparison (Branscombe et al., 2012), or reinterpreting the valence of the condition (Riley et al., 2009).

However, as argued throughout this thesis, for many HPC disorders, the viability of symptom improvement (i.e., upward mobility) is often highly contested, difficult to know in advance, may be a matter of degree, and tends to be revealed over time. For some, substantial symptom improvement may be rapid, complete and linear; for others it is gradual, partial and punctuated by setbacks (Klauer, Ferring, & Filipp, 1998). Consequently, a very real question that confronts many with mental illness is the question of whether, and to what

extent, symptom remission is possible (Howard, 2008). This is particularly true for a range of common psychological conditions such as Major Depression and anxiety disorders that can have varying degrees of severity, duration and intermittency. Lay understandings of these conditions (Angermeyer & Matschinger, 2003) are broadly consistent with clinical evidence indicating that while sustained remission is by no means certain, it is a more likely outcome compared to chronic conditions such as schizophrenia and bipolar disorder (Kessler et al., 2005) or developmental conditions such as autism.

But there is also evidence for considerable variation in lay beliefs about the extent to which conditions such as depression and anxiety are changeable (Haslam, 2002a). This perhaps reflects the difficult choices confronting individuals with such conditions (Howard, 2008). One option is for sufferers to engage in various evidence-based, and usually individualistic, psycho-therapies (such as cognitive therapy) in the hope of overcoming one's condition and thereby "leaving the group" (i.e. individual upward mobility). But doing so is problematic. Disidentifying with the group creates the risk that an individual will neither be accepted fully by other sufferers as truly unwell nor accepted by the healthy mainstream as "normal" (Howard, 2008), or perhaps only by "passing" as normal (Goffman, 1963). Moreover, to the extent that an individual engages in, for example, psycho-therapy that is seen to contribute to symptomatic recovery, this is likely to be perceived by others as implying that the condition is (in some sense) within that individual's control (Deacon & Baird, 2009). This again plays in to the logic of certain stigmatising views that those with mental illness are to blame for their condition (Deacon & Baird, 2009, Lebowitz, Puyn, & Ahn, 2014). In the face of these considerable threats (Howard, 2006), it is perhaps not surprising that sufferers may identify strongly with a disorder-based group with a shared, lay understanding that a condition is bio-genetically determined (i.e. fixed). Not only do such groups provide a sense of positive support, and the opportunity to engage in social creativity

strategies (Jetten et al., 2011), but by validating the biogenetic basis of a condition this helps to challenge stigmatising outgroup beliefs that people with conditions like depression or anxiety are blameworthy (Lebowitz, 2014).

However, as mentioned earlier, this comes at the considerable cost of inducing prognostic pessimism among sufferers (Kvaale et al., 2013). More troublingly, such a strategy may represent a significant barrier to engaging in treatment. This is particularly problematic given the clinical maintaining factors that underlie depressive and anxiety disorders (Dozois & Westra, 2005). For these conditions, the belief that conditions are unchangeable, and the subsequent avoidance of anxiety provoking stimuli (Abramson et al., 1989), is precisely what serves as a clinical maintaining factor. In other words, to the extent Disorder Identity groups promote a biogenetic theoretical understanding of a condition, sufferers are likely to develop a “fixed” rather than a “malleable” view of their emotional state that maintains, if not exacerbates, their clinical presentation (Kneeland et al., 2016). In terms of the dynamic processes described within the HPC framework, such a Disorder Identity, with a negative fixed group norm, would serve as a negatively heterostatic and positively homeodynamic mechanism.

In sum, on the evidence reviewed so far, the viability of adopting a Disorder Identity Adaption Strategy involving collective upward mobility appears problematic. Research examining implicit theories and mental health has focused on the negative clinical consequences of essentialist theories (Lebowitz, 2014) and fixed theories (Kneeland et al., 2016) and it is often noted that these beliefs can be promoted by Disorder Identity groups (Howard, 2008; Corrigan et al., 2009). This perhaps reflects a tacit assumption underpinning implicit theory research in particular that group-based perception is tied to a fixed, rather than a malleable, mindset (Levy et al., 2001), and so to the extent learning a more malleable mindset is possible then this requires more individuated perception rather than relying on

group-based stereotypes (Levy, Stroessner, & Dweck, 1998). Consequently, the possibility of adopting a collective mobility strategy, as outlined here, is seldom considered within this literature.

Despite this, as discussed in Chapter 3, there is evidence suggesting that the relationship between essentialist beliefs, including lay theories, and group perception is complex (Haslam et al., 2006). Specifically, there is empirical evidence that groups may flexibly switch between holding more or less essentialist beliefs—not as a function of group-based self-perception but as a function of the strategic goals of the group and the social context (Morton et al., 2009). This suggests that group-based self-perception and malleable beliefs about ingroup members, implied by the proposed collective mobility strategy, can indeed coexist, at least to a degree.

This evidence is broadly consistent with the SCT view that group based self-stereotyping is an inherently flexible, and adaptive, meaning-making process (McGarty et al., 2002). Despite this, there are important theoretical ambiguities within SIT about the possibility of collective mobility as proposed here. Specifically, as mentioned earlier, group-boundary permeability has typically been conceptualised as an externally determined socio-structural variable that influences the likelihood that low status group members will adopt *either* an individual upward mobility strategy to join the high status group *or* the various group-based strategies of social creativity and social competition. Although Tajfel and Turner (1979) stress that these various responses to a group-based social hierarchy occur on a continuum, in fact these strategies have mostly been studied empirically as categorically distinct (Ellemers et al., 1988). Very little SIT research has investigated circumstances where boundary permeability is ambiguous, and where boundary permeability is itself understood as a (potentially) modifiable characteristic of group members. However, it is often precisely

these circumstances that confront people with mental disorders such as depression and anxiety.

It is also important to explain in more detail why these circumstances are theorised to be particularly germane to adopting a *collective* mobility strategy and why this strategy is distinct from the more well-documented SIT strategies of individual mobility and social change. A social change strategy is predicted to occur where inter-group status relations are perceived to be illegitimate. The behaviours associated with this strategy are directed towards changing the socio-structural features of the environment (for example, by demanding that laws be changed or resources allocated more fairly to ingroup members). By contrast, collective mobility as outlined here is different in two key respects. First, it is not dependent on perceiving illegitimate status relations; and second, the focal behaviours are about changing the self rather than the social context.

Collective mobility is also distinct from individual upward mobility. Individual upward mobility implies that my personal malleability beliefs are primarily informed by my personal experience of setbacks and incremental progress, as well as expert opinion on the prognosis for my condition and wider ideologies about the viability of individual self-improvement. By contrast, collective mobility, as outlined here, suggests that fellow group members may provide a uniquely valuable epistemic and motivational resource for informing the possibility and desirability of change. Rather than relying on my personal experience of setback and failure, fellow sufferers provide evidence about how, whether and to what extent *people like us* can change and so, by implication, evidence about whether *I too* can change.

Collective upward mobility, then, appears to be a distinctive, yet under-theorised, response to being a member of a low status group. The task of this short chapter, therefore, is to develop a series of empirical propositions about the causes and consequences of adopting such a collective upward mobility strategy. For brevity, I will refer to these predicted

consequences of adopting this Disorder Identity Adaption Strategy as the “collective mobility thesis”.

The collective mobility thesis is predicted to be operative under circumstances in which Disorder Identity group members perceive that a condition is negatively valenced and malleable. These circumstances are likely to obtain for conditions for which sufferers experience a level of control over their symptomatology as a function of their beliefs about the nature of their condition. As discussed, anxiety disorders are a good example of such disorders. Motivational processes and lay theories of emotion are thought to play an important role in the aetiology and maintenance of anxiety (Dozois & Westra, 2005; Kneeland et al., 2016). Specifically, clinical research suggests that individuals who hold a more malleable view of their anxiety are likely to display better emotional self-regulation (Kneeland et al., 2016) and to be more likely to display approach motivation, rather than avoidance, in the face of anxiety provoking situations (Beer, 2002; Schroder et al., 2016). Over the longer term, a more malleable mindset about anxiety is therefore thought to help drive reductions in anxiety symptomatology.

The first priority, then, is to develop empirical predictions about the social psychological origins of personal malleability beliefs in the context of the experience of conditions such as anxiety. As discussed, there are two, conceptually distinct accounts of the origins of malleability beliefs corresponding to the collective mobility thesis and the individual upward mobility thesis that need to be compared. As I will explain shortly, these two accounts are not necessarily mutually exclusive, but they do generate quite different empirical predictions.

The individual mobility thesis is arguably a corollary of the claim that group based perception is linked to a more fixed mindset (Levy et al., 2001), and so individuated perception is associated with holding more malleable beliefs. Similarly, within SIT, upward

mobility is typically understood as an individualistic, rather than a collective, response to being a member of a low status group (Tajfel & Turner, 1979) and can be equated with holding more malleable beliefs about a condition. Moreover, as discussed earlier in this chapter, research by Ellemers et al. (1997) has shown that, all things being equal, low identifiers (with respect to a low status group) were more likely to seek to adopt an upward mobility strategy compared to high identifiers. Because theoretically, to the extent one identifies less as a group member, the more likely one's personal identity is salient (Onorato & Turner, 2004), this implies:

Empirical Proposition 1 (consistent with the individual mobility thesis): To the extent individuals' personal identities are made salient, they will, on average and all else being equal, perceive their own condition to be more malleable.

By contrast, an alternative account is that upward mobility is a collective, rather than an individuating process. The claim that has been developed in Chapter 5 is that collective mobility is a kind of Disorder Identity Adaption Strategy and so, by definition (see Chapter 2), is theorised to become operative to the extent a person self-categorizes in terms of a particular Disorder Identity group. Under these circumstances, group members are predicted to act in line with the normative understanding about what a disorder category means. Specifically, the collective mobility thesis implies that an individual's *personal* beliefs about the malleability of his or her own condition (e.g. "my anxiety can change") will be socially influenced by the malleability beliefs of fellow disorder-identity group members (e.g. "our anxiety can change"). This is predicted to occur because, as outlined in Chapter 4, shared social category membership provides the basis for social influence (Turner, 1991). This implies:

Empirical Proposition 2 (consistent with the collective malleability thesis):

To the extent individuals' Disorder Identities are made salient, those individuals will be influenced by fellow Disorder Identity group members to perceive their own condition to be either relatively fixed or malleable, consistent with the group norm.

Proposition 1 and 2 are not mutually exclusive. A main effect of personal identity salience on personal malleability beliefs would provide evidence for 1 (individual mobility), whereas a malleable group norm by Disorder Identity salience interaction effect on personal malleability beliefs would provide evidence consistent with 2 (collective mobility).

Since identity salience is partly determined by long term identification with a disorder group (McGarty & Grace, 1999), Empirical Propositions 1 and 2 imply the following Corollaries:

Corollary 1: (consistent with the individual mobility thesis, and Ellemers et al. (1997)) Those who identify less strongly with a Disorder Identity group (i.e. low identifiers) will perceive their own condition to be more malleable compared to those who identify highly with the group.

Corollary 2: (consistent with the collective mobility thesis) Those who identify highly with a Disorder Identity group will be influenced by fellow disorder group members to perceive their own condition as either relatively fixed or malleable, consistent with the group norm.

The collective mobility thesis further implies that Disorder Identities provide a basis for social influence of personal malleability beliefs not merely because of shared similarities between group members. In fact, the evidence that "mere similarity", on any dimension, provides a basis for social modelling of self-efficacy is weak and inconsistent (Usher & Pajares, 2008). Instead, the self-categorization process, and therefore the social influence

process, is further constrained by the theoretical relevance of a particular category. Because of this, fellow Disorder Identifiers are uniquely informative about the malleability of my condition in a way that ingroup members, from myriad other social groups, are not. This implies:

Empirical Proposition 3: Those who identify highly with a Disorder Identity group will be influenced by fellow disorder group members endorsing malleability beliefs, but will not be similarly influenced by fellow ingroup members from theoretically irrelevant groups endorsing the same malleability beliefs about the condition.

The Empirical Propositions and Corollaries developed so far, pertain to the social-psychological origins of personal malleability beliefs. However, the broader Disorder Identity Analysis implies not just that identity processes influence people's beliefs about their disorder but that this will ultimately influence how they experience symptoms (Levine, 1999). Specifically, it is predicted that the Disorder Identity Adaption Strategy of collective mobility should attenuate symptoms (i.e. be negatively homeodynamic) and facilitate the transition of group members from a disordered state to a non-disordered state (i.e. be positively heterodynamic) over time. This implies:

Empirical Proposition 4: to the extent individuals identify with a Disorder Identity group, and that group holds a malleable group norm, then it is predicted those individuals will display reductions in symptomatology over time.

Finally, the conjunction of Corollary 2, with extant evidence that personal malleability beliefs about disorder predict reductions in disorder symptomatology (Beer, 2002), and Proposition 3 imply:

Empirical Proposition 5: The interaction between a malleable group norm and disorder identification will predict more malleable personal beliefs about the disorder which will, in turn, positively mediate the relationship between the malleable group norm and lower symptomatology (i.e. conditional mediation, Hayes, 2012).

The upcoming empirical chapters aim to test the collective mobility thesis and aspects of the individual mobility thesis by testing hypotheses derived from the empirical propositions and corollaries outlined above. Broadly, the propositions and corollaries can be grouped into two sets. The first set (Empirical Propositions 1-3 and Corollaries 1-2) all pertain to the processes determining personal malleability beliefs. This set of propositions will be examined over three studies, set out within Chapter 6. The second set (Empirical Propositions 4-5) will be examined in Study 4 (Chapter 7) and Study 5 (Chapter 8). This second set pertains to the symptomatic and behavioural consequences of adopting a collective mobility strategy.

These propositions pose a number of methodological and empirical challenges. Among them, that most of the propositions entail testing two-way interaction effects within a putatively “disordered” population. Specifically, as mentioned earlier, the collective mobility thesis is likely to be particularly relevant to conditions such as anxiety disorders. However, given the primarily theoretical aim of the proposed investigation, the sample sizes required to test interaction effects in a clinically symptomatic population, and the ethical risks of manipulating beliefs about the malleability of a putative disorder, it seemed advisable to test these propositions using an analogue anxiety condition. Faced with similar challenges, experimental researchers have used a range of analogue disorders, including subclinical depression (Vredenburg, Flett, & Krames, 1993), shyness (Beer et al, 2002), and speech anxiety (Kneeland, Nolen-Hoeksema, Dovidio, & Gruber, 2016), to evaluate similar, basic

social-psychological processes in a clinical domain. If the proposed theoretical processes can be demonstrated in such analogue groups, then it lends credibility to the argument that such effects may apply to clinical disorders.

For reasons that will be outlined in the next chapters, the empirical studies reported here test the collective mobility thesis in the context of the experience of the subclinical conditions of shyness and speech anxiety. However, neither shyness, nor speech anxiety, have well established or validated normative criteria for defining a discrete “symptomatic” population. Where possible, theoretically relevant populations were identified by pre-screening participants whose symptomatology exceeded previously identified normative cut-offs.

Finally, the central theoretical construct of “malleability” has been defined quite broadly up until now. This was done deliberately to help integrate very diverse theoretical literatures, all of which invoke conceptually related, but nevertheless distinct, understandings of whether an individual perceives some property to be malleable. Boundary permeability as used within clinical applications of SIT (Jetten et al., 2011), change expectancy within the clinical literature (Dozois & Westra, 2005), lay incremental and entity theories within the implicit theory literature (Dweck et al., 1995), and intentionality within certain accounts of attribution (Haslam, 2005; Malle, 1999), all attempt (in different ways) to capture the extent to which an individual perceives that some personal characteristic can either vary as a function of effort and accumulated experience or is immutable and beyond any volitional control (see also Haslam, Bastian, Bain, & Kashima, 2006).

For the purposes of the current empirical program, this construct will be operationalised by using the implicit beliefs measure developed by Beer (2002) and Dweck (1995), given its demonstrated utility in predicting avoidance and approach in clinical contexts (Kneeland et al., 2016a). Of course this leaves open the question for future research

to determine whether this is the most appropriate way to operationalise the sense of malleability developed here. It also leaves open whether this operationalisation collapses across more fine-grained conceptual distinctions that might predict additional variance in the Disorder Identity dynamics theorised above.

Chapter 6: Identity Processes as Determinants of Individual Malleability Beliefs

In this chapter I investigate how Disorder Identity processes influence people's personal malleability beliefs about disorder. In doing so, across three experiments I test hypotheses derived from Empirical Propositions 1-3 (and their corollaries) in the context of people who experience significant symptoms of the analogue "disorder" of shyness. Briefly again, these Empirical Propositions are, firstly, that individuals will, consistent with the individual mobility thesis, perceive their own condition as relatively malleable to the extent their personal identity is salient (Empirical Proposition 1) or they are low Disorder Identity group identifiers (Corollary 1). However, consistent with the collective mobility thesis, individuals will also be socially influenced by fellow disorder sufferers to perceive their own condition as relatively fixed or malleable, consistent with the group norm, to the extent their Disorder Identity is salient (Empirical Proposition 2) and to the extent they identify with a disorder category (Corollary 2). And finally, individuals' personal beliefs about the malleability of their condition will be influenced by fellow disorder sufferers' opinions about the malleability of a condition to the extent they identify with the disorder. However, individuals will not be similarly influenced by the opinions about the disorder held by theoretically unrelated groups to which they identify (Empirical Proposition 3). Because these propositions and corollaries are all very closely related, in the sense that they all examine the social determinants of individual malleability beliefs, they are all examined within a single chapter.

In the first part of this chapter I briefly summarise the clinical features of shyness and outline why shyness provides an ideal Disorder Identity analogue condition for testing the proposed identity processes. I then reiterate the theoretical basis for testing Empirical Propositions 1-3 (and their corollaries) in terms of the basic tension between the individual mobility thesis and the collective mobility thesis. I then report the findings from the three

experimental studies and conclude by summarising the broader implications of these findings for the collective mobility thesis and the individual mobility thesis.

As mentioned in Chapter 5, for practical and ethical reasons it was not appropriate to test these proposed identity processes in the context of individuals experiencing clinically diagnosable disorders. The challenge, then, was to develop a suitable disorder analogue with which to test the social-psychological processes underlying the Disorder Identity Adaption Strategy of collective mobility. Ideally, such an analogue condition would display a number of features. Firstly, if such a condition is to be of clinical relevance, then ideally it should represent a less severe form of a clinically diagnosable condition. Secondly, the condition should ideally be sufficiently prevalent (for the pragmatic reason of needing to access adequate sample sizes) and there should be a degree of lay awareness about the condition. This latter point is important because, according to the broader Disorder Identity Analysis, it is in these circumstances that Disorder Identity-based mechanisms should be operative. And finally, in order to test the proposed collective mobility thesis specifically, it is important that the condition is generally perceived to be negatively valenced but that there is variability in lay perceptions about the malleability of the condition.

By these criteria, the sub-clinical condition of shyness seems well suited for the proposed analysis. Shyness is likely to be clinically relevant, given that a number of researchers have noted that social phobia is simply an extreme form of shyness (Henderson and Zimbardo, 1999). Turner et al. (1990) found that shyness shares many of the somatic (trembling, blushing), cognitive (fear of negative evaluation), and behavioural (distress and avoidance) symptoms of clinical social phobia. Rapee (1998) goes further in arguing that “many words and terms have been used to describe shyness, including social phobia, social anxiety, avoidant personality disorder...they all refer basically to the same thing” (p. xi). On

balance, then, the relatively limited corpus of research on shyness suggests that it is likely to be at least a plausible analogue condition for understanding Disorder Identity processes.

Moreover, as a psychological construct, shyness seems to have more than a superficial resemblance to many looping kind disorders (Hacking, 1999). Prevalence rates of shyness display marked cross-cultural and longitudinal variation (Henderson, Zimbardo, & Carducci 2010). For example, rates of shyness in the United States appear to have increased from about 40 percent (+/-3%) in the 1970s to 48 per cent (+/-2%) by the late '90s (Carducci & Golant, 2000), with more recent estimates suggesting a prevalence in college samples as high as 60 per cent (Henderson & Zimbardo, 2001). Some argue that this increase in prevalence reflects the increasing medicalising of normal human variation (Lane, 2006). However, others note that the condition can cause significant distress and impairment (Dalrymple & Zimmerman, 2013) and that most people report a desire to overcome their shyness (Henderson & Zimbardo, 2001). Indeed, since Zimbardo, Pilkonis, and Norwood's (1975) popular article titled the "Social Disease Called Shyness", the condition has been the subject of other popular articles (Holmes, 2014), as well as manualised psychological treatment programs aiming to "treat" shyness (Henderson & Zimbardo, 2001, Titov, Andrews, Schwencke, Drobny, & Einstein, 2008). There is evidence to suggest, then, that the condition is relatively prevalent, that there is some lay awareness about it as a phenomenon, and that the condition is widely considered to be negatively valenced.

Finally, both academic and lay theorising about the condition suggests divergent views about the extent to which the condition is malleable. Haslam, Bastian, and Bissett (2004) have measured people's lay beliefs about the essentialist basis of various personality characteristics and found that shyness was perceived to be close to the mid-point of their essentialism scale. Although the distribution of these judgements was not reported, the finding suggests that it is unlikely to be highly skewed, and so is potentially more amenable

to experimental influence than other characteristics for which people have strong pre-existing views about the nature of the condition. Consistent with this, the condition has provided the basis for previous experimental research examining the impacts of lay theorising about the malleability of shyness (Beer, 2002; Valentiner, Mounts, Durik, & Gier-Lonsway, 2011). In sum, for the reasons outlined above, the condition seemed well suited to assessing the collective mobility thesis.

To reiterate the arguments I presented in Chapter 5, one of the key elements of the collective mobility thesis is that fellow Disorder Identity group members are a powerful epistemic resource because they provide *valid* experience for informing individual group members' personal beliefs regarding the malleability of their shared condition. This claim raises a deeper question about where people's beliefs about the malleability of disorders, and essentialist beliefs more generally, come from (Haslam et al., 2006; Haslam & Ernst, 2002; Medin & Ortony, 1989). Specifically, to what extent are malleability beliefs themselves malleable and, if they are, *how* is it that an individual comes to believe that his or her personal experience of shyness can change?

A large body of research, touched on in Chapter 3, emphasises the evolutionary (Atran, 1998) and developmental origins of psychological essentialism (Gelman, 2004;. Gelman & Markman, 1986; Gelman & Wellman, 1991). Importantly, much of the research in this tradition has demonstrated how both children and adults defer to experts when it comes to categorizing and characterising essences (Gelman, 2009; Kalish, 1998; Malt, 1990). Similarly, research examining implicit theories initially suggested that such theories are individual difference variables that (perhaps ironically) are relatively stable and chronically held (Dweck et al., 1993) and may be largely shaped by early socialisation (Dweck et al., 1995; Mueller & Dweck, 1998) and individual experiences of success and failure (Bandura, 1997, Usher & Pajares, 2008).

A possible exception to this, mentioned in Chapter 3, is Bandura's social cognitive theory according to which "similar" others provide a basis for influencing self-efficacy beliefs, a process called social modelling (Bandura, 1977; 1986). Although self-efficacy beliefs, defined as the belief that one has the ability to succeed in a specific situation, are distinct from malleability beliefs, these constructs are clearly closely related. The empirical evidence for social modelling of self-efficacy beliefs has been quite mixed in part because of the challenges in conceptualising and operationalising the nature of "similarity" (Usher & Pajares, 2008). As Bandura (1997) points out, because people share myriad similarities, this poses problems for operationalising which similarities are likely to dominate.

A further line of research has shown that beliefs about essentialism and malleability can be changed through relatively brief inductions (Hong et al., 1999), as well as over the course of long-term training programs aiming to teach participants a "growth" mindset (Blackwell et al., 2007). While these insights have led to the development of influential applied interventions, again a prevailing assumption in much of this work is that implicit theories are best understood as individually held beliefs (Molden & Dweck, 2006) and that to the extent these beliefs can be changed, this is an outcome of expert influence. However, what constitutes "expert" knowledge is left relatively underexplored. It seems that a common, tacit, assumption in these literatures is that individuals seek out "objective authority" on the nature of reality and that only this informational influence leads to private acceptance that information is true (Deutsch & Gerard, 1955).

In contrast to this, other research emphasises the strategic dimension to the origins of broader essentialist beliefs (Yzerbyt, Corneille, & Estrada, 2001). In intergroup contexts, for example, research suggests that individuals are more likely to ascribe positive human essences to ingroup members than outgroup members (Leyens et al., 2001)—in part, according to the authors, to justify and explain the ingroup's superiority. There is also

research, mentioned in Chapter 3, showing how group members flexibly and selectively invoke essentialist theories to counter different kinds of discrimination (Morton, Hornsey, & Postmes, 2009; Morton & Postmes, 2009; Morton, Postmes, Haslam, & Hornsey, 2009). But proponents of Deutsch and Gerard's (1955) model of influence might wonder if these strategic shifts in beliefs about the malleability of a condition are instances of mere normative influence—generating relatively superficial conformity to normative beliefs in inter-group contexts that does not translate into deeper acceptance in private. The implication of this is that although individuals may purport to endorse the claim that shy people in general can change (or not) when in a particular intergroup context, this may not necessarily imply they will believe that they, as individuals, can change.

Finally, in Chapter 5, it was noted that although the Social Identity Approach suggests an alternative account of how people's malleability beliefs are socially influenced, there are also significant ambiguities about how this might play out. To reiterate, according to Social Identity Theory, when people perceive themselves to be in a low-status group they will strive to positively define themselves in different ways, depending on the features of the intergroup context. One of the main features that governs whether people are more likely to adopt a strategy of individual upward mobility (i.e. to strive to leave the shy group as an individual) or more group-based strategies of social creativity (e.g. "we are shy but considerate") is the perceived permeability of the intergroup boundary. Although boundary permeability is often construed as an obvious feature of the social context, for groups defined by internal attributes (Tajfel, 1978), boundary permeability can be understood as whether or not a condition is perceived to be malleable. However, previous research suggests that identification with a low status group is both a cause and a consequence of perceived boundary *impermeability* (Ellemers, Spears, & Doosje, 1997; Ellemers et al., 1988; Yzerbyt, Castano, Leyens, & Paladino, 2000). This suggests that disorder group identification is a barrier to believing in

the possibility of change, and that upward mobility is an inherently individuating process. Furthermore, in the context of Disorder Identities, a putative ingroup member who claims that it is desirable and easily possible for us to overcome our shyness may be unlikely to be perceived as “one of us” because she potentially exposes us to moral reproach if we subsequently fail to change (Weiner, 1985). This reflects a widely noted concern that fostering a strong sense of identification with a disorder category may strengthen people’s beliefs that a disorder is immutable and, conversely, that more individuated perception is linked to a more malleable mindset (Levy et al., 1998). For this reason, in Chapter 5, I developed Empirical Proposition 1, according to which, to the extent that an individual self-categorizes in terms of their personal identity, rather than their collective, Disorder Identity, then that individual is more likely to perceive his or her condition to be more malleable (rather than fixed).

By contrast, on the basis of Self-Categorization Theory it might be argued that fellow Disorder Identity group members are uniquely informative about the malleability of a person’s shyness, since shared social-psychological group memberships provide the basis for mutual social influence (Turner 1991). Crucially, it is not just any shared membership group that counts (c.f., Bandura, 1997). What matters is that fellow ingroup members share the same condition as me and so have a shared experience of what it is like to be shy. If these fellow Disorder Identity group members perceive that “our shyness can change”, then this will inform my belief that “*my* shyness can change”. It was this possibility that underpinned the claim that transitioning out of a Disorder Identity group could be understood as a group-based, rather than an individuating, process (i.e. the collective mobility thesis).

In the first two studies of this chapter, I therefore pit the social influence processes underlying the individual mobility thesis and collective mobility thesis against each other. To reiterate, these two theses correspond, respectively, to:

Empirical Proposition 1 (consistent with the individual mobility thesis): To the extent individuals' personal identities are made salient, they will perceive their own condition is more malleable.

Empirical Proposition 2 (consistent with the collective mobility thesis): To the extent individuals' Disorder Identities are made salient, those individuals will be influenced by fellow Disorder Identity group members to perceive their own condition to be either relatively fixed or malleable, consistent with the group norm.

Because the salience of a self-category is partly determined by a person's long term identification with a social group (McGarty & Grace, 1999), and because to the extent a person's social identity is salient this inhibits the salience of personal identity (Onorato & Turner, 2004), these propositions imply the following corollaries:

Corollary 1: (consistent with the individual mobility thesis, and Ellemers et al. (1997)) Those who display low levels of identification with a disorder group will perceive their own condition to be more malleable (rather than fixed) compared to those who identify highly with the group.

Corollary 2: (consistent with the collective mobility thesis) Those who identify highly with a Disorder Identity group will be influenced by fellow disorder group members to perceive their own condition as either relatively fixed or malleable, consistent with the group norm.

Finally, in Study 3, I test the further claim, implicit in the collective mobility thesis and the broader Disorder Identity analysis, that Disorder Identities are important because fellow sufferers provide validated information about whether my shyness can change. In other words, fellow sufferers are perceived as "experts by experience". By contrast, ingroup members from other groups that might be psychologically meaningful to me (e.g. gender,

national, or occupational groups) may hold shared beliefs about the nature of shyness-but these beliefs are less likely to inform my beliefs about the malleability of my shyness because these fellow “ingroup” members are not similar to me in a theoretically relevant way. Study 3 therefore sought to confirm this claim by assessing:

Empirical Proposition 3: Those who identify highly with a Disorder Identity group will be influenced by fellow disorder group members endorsing malleability beliefs, but will not be similarly influenced by fellow ingroup members from theoretically irrelevant groups endorsing the same malleability beliefs about the condition.

Study 1 introduction

Study 1 aimed to test Corollary 1 and Corollary 2 by experimentally manipulating the group norm and assessing how this, in interaction with measured social identification as shy, affects participants’ personal malleability beliefs. Corollary 1 can be equated with the hypothesis that high (low) levels of identification overall will be positively correlated with more fixed (malleable) personal beliefs about the nature of shyness (Hypothesis 1). Corollary 2 in the context of shyness as a disorder analogue, can be equated with the hypothesis that to the extent that people identify as shy they will be influenced by fellow shyness sufferers to perceive their own condition as either relatively fixed or malleable, consistent with the group’s normative beliefs about the malleability of shyness (Hypothesis 2).

Study 1 method

Participants and design

Thirty-nine eligible study participants (median age = 49.5 years; 19 males) completed the study. Potential participants were recruited through a web-based survey company, the Online Research Unit. Participants were invited to complete a computer based study aimed at “understanding the factors affecting shyness” in exchange for remuneration from the Online Research Unit. The study design manipulated the malleability of the group norm

within the group of shy people (malleability: high, low). Participants' level of identification with the group of shy people was a measured, predictor variable.

Materials and procedure

Participants were first asked "how shy are you", and asked to respond using a 1 ("not at all shy") to 5 ("extremely shy") Likert scale. This item has been demonstrated to be a valid and reliable screening measure for assessing shyness symptoms (Hopko, Stowell, Jones, Armento, & Cheek, 2005). The computer was programmed to screen those participants who rated their shyness at 3 or above ($N=39$). Participants who reported levels of shyness below the midpoint of the scale were thanked for their time and did not answer any further questions ($N=49$). This screening procedure ensured that only those who displayed symptoms of shyness were recruited.

Eligible study participants were then asked a set of more in-depth questions assessing their shyness symptoms, using the 13 item Revised Cheek and Buss Shyness Scale (RCBS), including items like "I am socially somewhat awkward". Participants responded to this using a five point scale from 1 ("never") to 5 ("always"). Participants then rated their individual fixed beliefs (Personal fixed belief Time 1), assessing the perceived malleability of their own shyness, using five items adapted from Beer's (2002) study of implicit beliefs and shyness (e.g., "I have a certain amount of shyness and it's something I can't do much about"). The items used for this measure are shown in Table 7. This scale and all other subsequent continuous rating scales were measured using a one ("strongly disagree") to seven ("strongly agree") point Likert scale. Participants were then asked to fill in a set of seven questions assessing their level of identification with the group of shy people, based on selected items from the social identity measure developed by Leach et al. (2008) and Postmes et al. (2012). The items used are shown in Table 7. These items were selected on the basis that they all formed plausible, natural language statements in the context of assessing identification as a

speech anxiety sufferer. To measure the perceived valence of shyness, participants were then asked a single item assessing their agreement with the statement “the benefits of being shy outweigh any disadvantages”.

After completing these introductory questions, participants were introduced to the main group norm manipulation. Participants were presented with a “word cloud” with 18 words of various sizes either relating to the relatively fixed nature of shyness (in the “fixed norm” condition; e.g. “can’t change”, “stuck”, “biological”), or the relatively malleable nature of shyness (in the “malleable norm” condition; e.g. “can change”, “transition”, “can overcome”, “learn”). Both conditions included various words relating to shyness and anxiety in general (e.g., “shy”, “tremble”). The word cloud was based on actual words used within different online shyness discussion forums. It was explained to participants that the researchers “used special software to analyse selected keywords from an online discussion forum of people who experience shyness. The larger the words the more often it is used in the shyness discussion forum”. Pictures of the fixed condition word cloud and the malleable condition word cloud are shown in Figure 8a and 8b respectively.

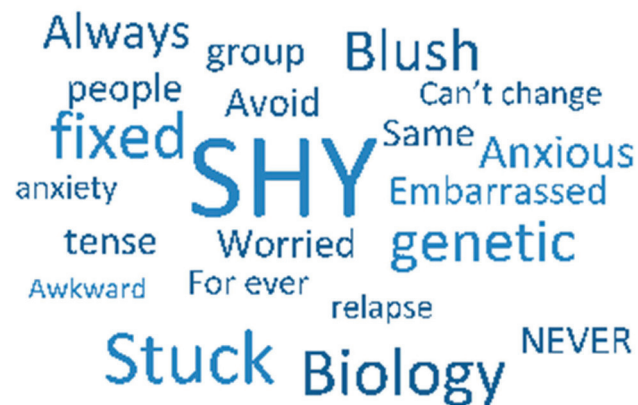


Figure 8a: "Fixed" word cloud



Figure 8b: "Malleable" word cloud

As a manipulation check, participants then answered five items measuring the perceived group norm of malleability, using modified items from Beer (2002) used to measure Time 1 personal fixed beliefs (e.g. "Shy people have a certain level of shyness, and it is something that they can't do much about").

Participants then rated their personal fixed beliefs at Time 2, using the same items used to assess personal fixed beliefs at Time 1.

After completing these questions, all participants were asked their sex and age and debriefed about the nature of shyness. All participants were reminded that shyness is very common, that it is not a clinical disorder, and provided evidence-based tips for managing shyness symptoms.

Table 7: *Items Measuring Social Identification and Individual Fixed Beliefs*

Measure	Item
Social Identification (Leach et al., 2008, Postmes et al., 2012)	I have a lot in common with the average shy person.
	I am similar to the average shy person.
	I feel a bond with shy people.
	Being shy is an important part of how I see myself.
	Shy people are very similar to each other.
	Shy people have a lot in common with each other.
	I identify with people who are shy. (from Postmes et al., 2012)
Individual Fixed Beliefs (Beer, 2002)	I have a certain level of shyness and it's something I can't do much about.
	My shyness is something about me that can't change very much.
	I can change aspects of my shyness if I want to. (Reverse Scored)
	How shy I am as a person is something that changes through my life. (Reverse Scored)
	My shyness is not fixed but changes over time. (Reverse Scored)

Study 1 results

Preliminary analyses and manipulation checks

The 13 items from the RCBS (Cronbach's $\alpha = .82$), the five items from the fixed beliefs about shyness at time 1 (Cronbach's $\alpha = .85$) and time 2 (Cronbach's $\alpha = .84$), and seven social identification items (Cronbach's $\alpha = .77$) all formed reliable scales. A series of principal components analyses were conducted on the sets of items within each scale and this revealed that the items within each scale all loaded on a single principle component with an eigenvalue greater than 1. This confirmed that the scales were unidimensional. The RCBS score was calculated by reverse scoring negatively worded items and then summing across all items. The mean RCBS score in the sample ($M = 40.91$, $SD = 7.31$) was more than one standard deviation higher than the mean reported in norms developed by Hopko et al. (2005), who report a mean of 31.1 ($SD = 8.2$) in population samples from the United States. This indicated that the screening question successfully selected participants with above average levels of shyness.

Before conducting the main analysis, preliminary analyses were conducted to test whether the shyness paradigm was interpreted as intended. Firstly, the valence of shyness was assessed by testing whether the mean ($M = 3.57$, $SD = 1.30$) was significantly below the mid-point of the seven point scale. This analysis revealed that, on average, participants disagreed that the benefits of being shy outweighed any disadvantages (mean difference = $-.43$, $t(38) = -2.14$, $p = .04$), suggesting that, as hoped, the subclinical category "shyness" was perceived to be more negatively than positively valenced.

Next, the impact of the group norm manipulation was assessed. The fixed group norm measure was significantly higher in the fixed norm condition ($M = 4.46$, $SD = .76$) than the malleable norm condition ($M = 3.80$, $SD = .59$, $t(37) = -2.97$, $p < .01$), indicating that the "word cloud" manipulation was successful.

Main analysis

The means, standard deviations and bivariate correlations are shown in Table 5. This shows that personal fixed beliefs (at both Time 1 and Time 2) were positively correlated with social identification with shyness, and so, by implication that low identification is associated with more malleable beliefs. This finding is consistent with Hypothesis 1 and the individual mobility thesis.

Table 8: *Study 1 means, standard deviations and bivariate correlations between main study variables.*

Measure	M	SD	1.	2.	3.	4.
1. RCBS	40.91	7.31	--	.71**	.56**	.61**
2. Personal Fixed Belief (T1)	4.17	.98		--	.44***	.83***
3. Shy identification	3.95	.97			--	.32**
4. Personal Fixed Belief (T2)	2.94	.76				--

** Correlation significant at $p < .01$ (two-tailed).

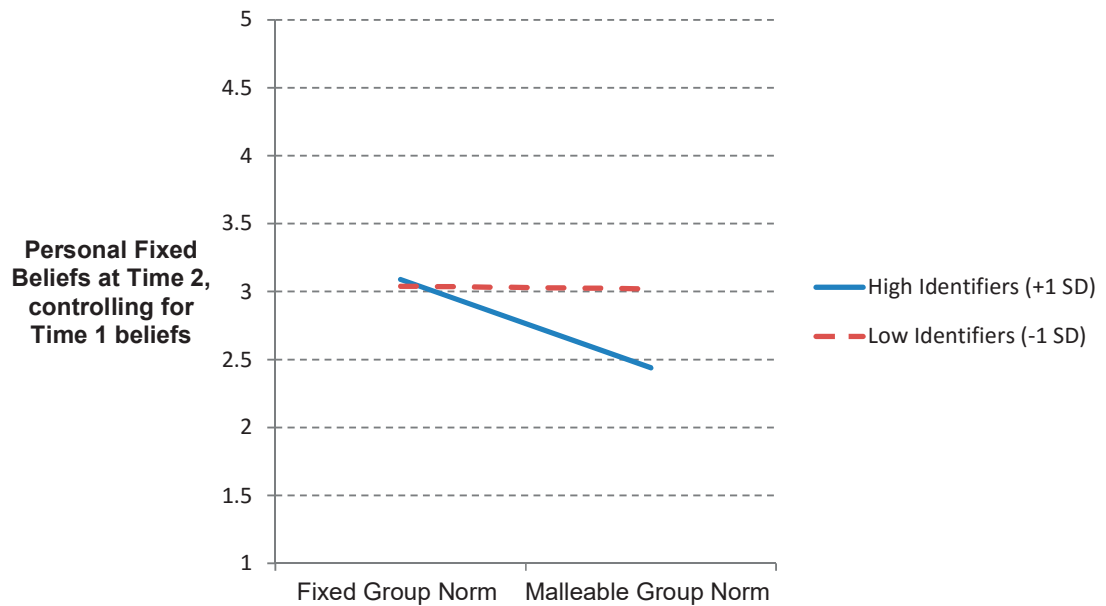


Figure 9: Significant two-way interaction between shy group identification and manipulated malleability norm on personal fixed beliefs at Time 2 (controlling for Time 1 fixed beliefs). Only highly identified shy group members were influenced by the malleable norm.

To test Hypothesis 2 (that high identifiers would be socially influenced by fellow shy group members to perceive their personal shyness was more fixed or malleable, consistent with the group norm) a 2 (group norm: fixed, malleable) x 2 (identification: higher, lower) analysis of covariance was run. Group norm was a manipulated independent variable, identification was a continuous (mean-centred) independent variable, Time 1 personal fixed belief was included as a continuous covariate, and Time 2 personal fixed belief was the dependent variable. The hypothesised two-way interaction between group norm and identification was significant ($F(1,38) = 4.91, p = .03$) and there was also a main effect for the group norm manipulation ($F(1,38) = 7.76, p = .01$), and the covariate ($F(1,38) = 91.80, p < .001$). The interaction effect was probed by analysing the main effect at one standard deviation above (high identifiers) and below (low identifiers) the mean level of social

identification. Controlling for Time 1 individual fixed beliefs, among high identifiers, Time 2 individual fixed beliefs were significantly higher in the fixed norm condition ($M = 3.09$) than in the malleable norm condition ($M = 2.44$, $t(38) = 3.37$, $p < .05$). By contrast, for low identifiers, there was no significant difference between the fixed ($M = 3.04$) and malleable norm conditions ($M = 3.02$, $t(38) = .14$, ns). These findings, confirming Hypothesis 2, are shown in Figure 9.

Study 1 discussion

Study 1 confirmed that high, but not low, identifiers were influenced by fellow shy people about the malleability of their personal shyness (Hypothesis 2), but that overall, low identification with the shy group is associated with more malleable beliefs (Hypothesis 1). This study was also important because it demonstrated the viability of shyness as a paradigm for testing group processes in an analogue disorder. In particular, the “word cloud” manipulation was perceived as intended, and participants consistently perceived that shyness was a negatively valenced condition.

Overall, these findings provided support for both Corollary 1 (and the individual mobility thesis) and Corollary 2 (and the collective mobility thesis). Specifically, although identifying with the shy group is correlated with more fixed beliefs about the nature of shyness, identification with the shy group appeared to provide a basis for inferring that the group’s normative understanding about the nature of shyness applies to the personal self. In other words, for high identifiers, knowing that shy people, “like me”, believe they can change provides a basis for believing that *I* can change. Interestingly, the pattern of means suggested that the two-way interaction was driven by high identifiers being influenced in the malleable norm condition that their shyness was significantly more malleable—rather than because high identifiers perceived that their personal shyness was more fixed when presented with a fixed norm. This is perhaps somewhat surprising. Given that identification appeared to be strongly

correlated with prior personal fixed beliefs, it might have been assumed that highly identified participants understood the category as a whole to be relatively fixed. When presented with a group norm claiming the group had relatively malleable views about the nature of shyness, this arguably may have undermined the normative fit of the shared self-category and therefore attenuated the strength of influence observed. Because shy group identification was a measured, rather than a manipulated, variable, it was not possible on the basis of the results of this study to confidently rule out that other prior beliefs may have altered the pattern of influence observed. Therefore, in Study 2, it was hoped that by experimentally manipulating identity salience, this would provide a clearer insight into whether the influence process was primarily driven by shared self-categorization, rather than prior beliefs that were associated with shy group identifiers.

Study 2 introduction

Study 2 sought to build on the Study 1 finding that high identifiers (compared to low identifiers) were more strongly influenced by fellow shyness sufferers through an experimental manipulation of both the group norm and identity salience. By either making participants' personal identities salient or their shy group identities salient, it was hoped that it would be possible to compare Empirical Proposition 1 (consistent with the individual malleability thesis) and Empirical Proposition 2 (consistent with the collective malleability thesis). Moreover, by experimentally manipulating relative group identity salience, the paradigm would provide more robust *causal* evidence for the proposition that self-categorization processes were driving the effects observed in Study 2.

There were two hypotheses in Experiment 2. In the personal identity salience condition, it was hypothesised that there would be a main effect of Personal Identity salience on personal malleability beliefs (Hypothesis 1, consistent with Empirical Proposition 1 and the individual upward mobility thesis). By contrast, Hypothesis 2 was that, after controlling

for participants' Time 1 personal beliefs about the malleability of their shyness, participants would be influenced by the group norm, but only in the shy group salience condition (i.e., consistent with Empirical Proposition 2 and the collective mobility thesis). In other words, it was predicted there would be a significant main effect of low identity salience on individual malleability beliefs and a significant two-way interaction between the group norm and shy salience on individual malleability beliefs.

Study 2 method

The methodology for Study 2 shared some similarities to Study 1. Substantive differences in the methodology are noted below.

Participants and design

One hundred and eleven participants were eligible to participate. Of these, 98 were included in the final analysis (median age = 55 years; 49 males), after excluding 13 individuals who failed the manipulation checks (described shortly). Potential participants were again recruited through a web-based survey company, the Online Research Unit, as described in Study 2.1. The study design manipulated the malleability of the group norm within the group of shy people (malleability: high, low), and also manipulated identity salience (salience: individual, shy group member).

Materials and procedure

Throughout the initial phase of the study, attempts were made to minimise the focus on shyness, so as to reduce the risk that the study procedures undermined the individual identity salience manipulation. Participants were invited to participate in a study investigating "participants' perceptions of themselves and others". Participants were initially asked to answer a series of questions assessing their eligibility for the study. Participants were first asked their gender and then asked the same "How shy are you" question that was asked in Study 1. Again, the computer was programmed to screen participants who rated

their shyness at 3 or above on this item. Participants who reported levels of shyness below the midpoint of the scale were thanked for their time and did not answer any further questions ($N = 129$).

Eligible study participants were then asked to answer a series of questions about “You and your relationships to other people”. This again included items from the 13 item Revised Cheek and Buss Shyness Scale (RCBS). Participants were then asked to answer the same five item measure assessing personal fixed (implicit theory) beliefs (Beer, 2002), prefaced with the introductory statement “Thinking about times in the past that you have felt shy...”.

Participants were then told that they had finished the first part of the survey and that they would now be asked a set of “different questions about how you see yourself and others”. Participants were asked a set of twelve filler questions about various personality attributes (e.g. “I like to tidy up”, “I am open minded”). It was hoped that this would serve as a conceptual break in the survey, and therefore minimise any carry-over effects from the first questions that focussed exclusively on shyness.

Participants were then introduced to the main identity salience manipulation using the “three things” manipulation (Haslam et al., 1999). Participants in the individual salience condition were told that their responses to the previous questions “indicate some of the things that make you a unique individual” and were then asked to write down three things that “you personally do relatively more often than most other people”, as well as things that they did “more rarely”, “more badly”, “differently”, and “better” than most other people.

In the shy identity salience condition, participants were told that their responses to the previous questions indicated “you do experience shyness” and were then asked to write down the four sets of three things. However, rather than being asked to write about their personal characteristics, they were asked about what “you and most other shy people” do compared to “non-shy people”. As a manipulation check, participants in both the shy salience and

individual salience conditions were asked if the questions above were asking participants to respond “as an individual”, “as a shy person”, or “as a tidy person”.

After completing the salience manipulation, participants were introduced to the main group norm manipulation, which used the same word clouds as used in Study 1. As a manipulation check, participants were asked after the word cloud to rate on a nine point bipolar continuum, whether the words used more often in the online forum related more to the relatively fixed nature of shyness or the relatively malleable nature of shyness. Participants then answered five items measuring the perceived group norm of malleability, again using modified items from Beer (2002). To reinforce the identity salience manipulation, participants in the shy salience condition were asked to rate their agreement with statements such as “we believe we have a certain amount of shyness and it is something we can’t do much about”, whereas in the individual salience condition, participants were asked to rate their agreement with slightly differently worded items, such as: “they believe they have a certain amount of shyness and it is something they can’t do much about”. It was hoped this use of the first person plural pronoun, “we” (as opposed to “they” in the personal identity salience condition), would emphasise that the people whose beliefs were reflected in the word-cloud shared a social category membership with participants.

Participants then rated their beliefs about the relative malleability of their own shyness, again adapting the five items from Beer’s (2002) study of shyness, but this time expressed in the first person (e.g., “I have a certain amount of shyness and it’s something I can’t do much about”). These items measured the main dependent variable, Time 2 personal fixed beliefs.

After completing these questions, all participants rated their level of identification with the group of shy people, using the same seven item scale used in Study 1. Finally, participants recorded their age and were debriefed about the nature of shyness. All

participants were reminded that shyness is very common, that it is not a clinical disorder, and provided evidence-based tips for managing shyness.

Study 2 results

Manipulation checks and preliminary analyses

Thirteen participants were excluded from the analysis because they failed the salience manipulation check assessing which group they had been part of. The final sample was 98. The impact of the group norm manipulation was then assessed. In the fixed norm condition, participants perceived that the words used in the word cloud related more to the relatively fixed nature of shyness ($M = 6.95$, $SD = 1.57$) than the malleable norm condition ($M = 4.02$, $SD = 1.98$, $t(96) = 8.31$, $p < .001$), indicating that the “word cloud” manipulation was successful.

The 13 items from the RCBS (Cronbach’s $\alpha = .83$), the five items from the personal fixed beliefs about shyness at Time 1 (Cronbach’s $\alpha = .86$) and Time 2 (Cronbach’s $\alpha = .87$), and seven social identification items (Cronbach’s $\alpha = .77$) all formed reliable scales. To assess whether each scale was unidimensional, the items within each scale were subject to a principal components analysis. All scales had single eigenvalues greater than 1, indicating that all of the scales were unidimensional. The means, standard deviations and bivariate correlations are shown in Table 9. The mean RCBS score was more than one standard deviation higher than the mean reported in norms developed by Hopko et al. (2005), indicating that the screening question successfully selected participants with above average levels of shyness.

Table 9: *Study 2, means, standard deviations, and bivariate correlations between main study variables.*

Measure	M	SD	1.	2.	3.	4.
1. Revised Cheek and Buss Shyness Scale (RCBS)	40.22	6.70	--	.58**	.28**	.45**
2. Personal Fixed Belief (T1)	3.86	1.20		--	.31**	.80**
3. Shy identification	4.20	.87			--	.32**
4. Personal Fixed Belief (T2)	3.73	1.12				--

** correlation is significant at the .01 level (2-tailed)

Study 2 main analysis

To test Hypothesis 1 and Hypothesis 2, a 2 (identity salience: personal identity, shy identity) x 2 (group norm: fixed, malleable) analysis of covariance was conducted, with Time 1 personal fixed beliefs as a continuous covariate, and Time 2 personal fixed beliefs as the dependent variable. This analysis revealed a significant two-way interaction between group norm and group salience ($F(1, 93) = 4.78, p = .03$), a main effect for the fixed group norm manipulation ($F(1,93) = 11.28, p < .01$), and the covariate ($F(1,93) = 204.78, p < .001$). There was no main effect of salience ($F(1,93) = .63, ns$), indicating, contrary to Hypothesis 1, that making participants' personal identity salient did not increase their personal malleability beliefs.

The interaction effect was decomposed by testing the effect of the group norm on Time 2 individual fixed beliefs, under individual and shy identity salience conditions, while controlling for Time 1 beliefs. Participants in the Personal Identity salience conditions rated their shyness as significantly more malleable under the malleable norm condition ($M=3.29$) than in the fixed norm condition ($M = 3.97$, $t(93) = 3.61$, $p < .001$). For the shy identity salience condition, participants did not differ significantly in the fixed norm condition ($M = 3.84$) compared to the malleable norm condition ($M = 3.74$, $t(93) = .54$, *ns*). Therefore, contrary to Hypothesis 2, there was no evidence that making the shy identity salient lead to more malleable or fixed beliefs, consistent with the group norm.

Since these findings did not confirm either hypothesis, a follow up analysis was conducted to assess whether these findings were further moderated by participants' measured social identification; this also served, in part, to assess whether the findings from Study 1 could be replicated. Since social identification in this study was measured after the main manipulation, a two-way analysis of variance was conducted assessing whether the two manipulated variables, and their interaction, influenced participants' level of social identification. This revealed no significant effects (all $p > .05$), indicating that social identification could be included in the analysis as a predictor variable.

A 2 (salience: high, low) x 2 (group norm: fixed, malleable) x 2 (measured, mean-centred, identification: higher, lower) analysis of covariance was conducted including Time 1 personal fixed beliefs as a covariate. This revealed a significant two-way interaction between the fixed group norm and group salience ($F(1,89) = 6.63$, $p = .01$), and a significant three-way interaction between fixed group norm, group salience and measured identification ($F(1,89) = 4.79$, $p = .03$).

To probe the three-way interaction, the interactive effects of norm and salience were analysed at one standard deviation above (high identifiers) and below the mean (low

identifiers) of measured identification. For high identifiers, under high identity salience conditions, there was a significant simple main effect of the group norm. High identifiers, under shy salience conditions, displayed significantly more malleable beliefs in the malleable norm condition ($M = 3.48$) compared to the fixed norm condition ($M = 4.01$, $t(89) = 2.08$, $p < .05$). The pattern of means was in a similar direction for high identifiers under individual salience conditions, but the difference was not significantly different ($t(89) = .86$, *ns*).

Low identifiers, under individual salience conditions, had significantly more fixed individual beliefs in the fixed norm condition ($M = 4.10$), compared to the malleable norm condition ($M = 2.99$, $t(89) = 4.44$, $p < .001$). Low identifiers, under shy salience conditions, displayed the opposite pattern, but the difference was not significantly different ($t(89) = -1.35$, *ns*).

Overall, the pattern of means in the follow-up analysis suggested two simple effects. Consistent with Corollary 2, and consistent with the results from Experiment 1, the malleable group norm influenced high identifiers under high shy salience conditions to display more malleable personal beliefs compared to when exposed to the fixed group norm. Secondly, and unexpectedly, when exposed to the malleable group norm, low identifiers, under individual salience conditions, showed more malleable personal beliefs compared to when exposed to the fixed group norm.

Follow-Up conditional mediation analysis

The foregoing analyses suggested that low identifiers under low group salience conditions and high identifiers, under high group salience conditions were affected by the word cloud manipulation. But what is not clear is whether both of these effects were driven by social influence. It would obviously conflict with Social Identity theorising, as well as extensive empirical research (see Haslam et al., 1999), if low identifiers, under low salience conditions, were subject to ingroup influence.

An alternative explanation is that low group identifiers, under low salience conditions, were, in fact, *differentiating* themselves from the group norm, rather than conforming to it. For example, it could be that these respondents were effectively communicating “they might think their shyness is malleable, but *my* shyness is even more malleable” (when presented with a malleable group norm). Consistent with this, it is notable that compared to the overall Time 2 personal fixed beliefs mean ($M = 3.73$), the effect appeared to be primarily driven by low identifiers, under low salience, malleable norm conditions, rating the fixedness of their shyness as being particularly low ($M = 2.99$). By contrast, low identifiers, under low salience, *fixed* norm conditions rated the fixedness of their shyness as only somewhat higher ($M = 4.10$) than the overall mean ($M = 3.73$). Importantly, a differentiation-based explanation would be theoretically consistent with Self-Categorization Theory. When individuals perceive themselves to share a salient social category membership, they perceive themselves more homogeneously and will accentuate the perceived similarities between group members (Turner et al., 1994). But conversely, when individuals do not share a salient group membership (which is particularly likely for low identifiers under low identity salience conditions) it would be expected that individuals will actively accentuate any differences between the self and salient “other” (in this case, the shy group).

However, to interrogate this possibility requires a measure of respondents’ *perceptions* of the group norm. This would provide a basis for assessing whether respondents were either conforming their personal malleability beliefs, to align with what they perceived the group’s beliefs to be (as in social influence) *or* were differentiating themselves from these group beliefs. One such indicator of respondents’ perceptions of the group norm is provided by the manipulation check asking respondents to rate on a continuum to what extent respondents perceived that the words expressed by the group of shy people in the word cloud

reflected the relatively fixed nature of shyness or the relatively malleable nature of shyness. Henceforth, this variable will be called “perception of the group norm”.

It is important to clarify the logic of a model aiming to test the alternative explanations for the two subgroups: high identifiers, under high salience conditions, and low identifiers, under low salience conditions. For each of these subgroups, the alternative explanations are whether the subgroup is: 1) conforming to the norm (consistent with social influence); or 2) differentiating from the norm. Conditional mediation analysis (Preacher, Rucker, & Hayes, 2007) is ideally suited to this sort of analysis, aiming to assess how the relationship between an independent variable (i.e. fixed norm manipulation) and a dependent variable (i.e. Time 2 fixed beliefs) can operate via different mechanisms (i.e. via a mediator, in this case the perception of the group norm) conditional on various moderating variables (i.e. participants’ identification as shy, and the salience of their shy identity). Such models allow the overall relationship between the independent variable and the dependent variable to be decomposed into an *indirect effect* (via the mediator) and a *direct effect*, while statistically detecting whether these two effects are operative only under particular conditions (or, in this case, for particular subgroups).

A model for testing the two alternative accounts, social influence or differentiation, is depicted in Figure 10, corresponding to Hayes’ (2012) Process model 19. The logic of this model is most straightforward for the case of social influence. Across all study participants, irrespective of identity salience or identification, it should be the case that the group norm manipulation will predict perceptions of the group norm (note, this is an assumption that will be tested shortly). However, perceptions of the group norm (as fixed) should only go on to predict personal Time 2 fixed beliefs for those subgroups where social influence is occurring. In this case, the interaction between identity salience and identification is predicted to moderate the relationship between perceptions of the group norm (the mediator) and

participants' Time 2 personal fixed beliefs (the dependent variable). An overall significant indirect path for a particular subgroup would be consistent with the claim that that subgroup's beliefs corresponded to their beliefs about what the group was saying, in line with the social influence explanation.

However, in the case where a subgroup is *differentiating* from the group norm (rather than conforming to it), the logic is importantly different. In this case, it would be expected that the word cloud manipulation will predict additional variance in Time 2 personal fixed beliefs *over and above* participants' perception of the group norm (i.e. while controlling for the indirect effect). This would correspond to a direct effect of the norm manipulation, but, again, conditional on participants' identity salience and level of group identification. A significant *direct* effect for a particular subgroup would be consistent with subgroup members differentiating themselves from what they perceived the group normative beliefs to be.

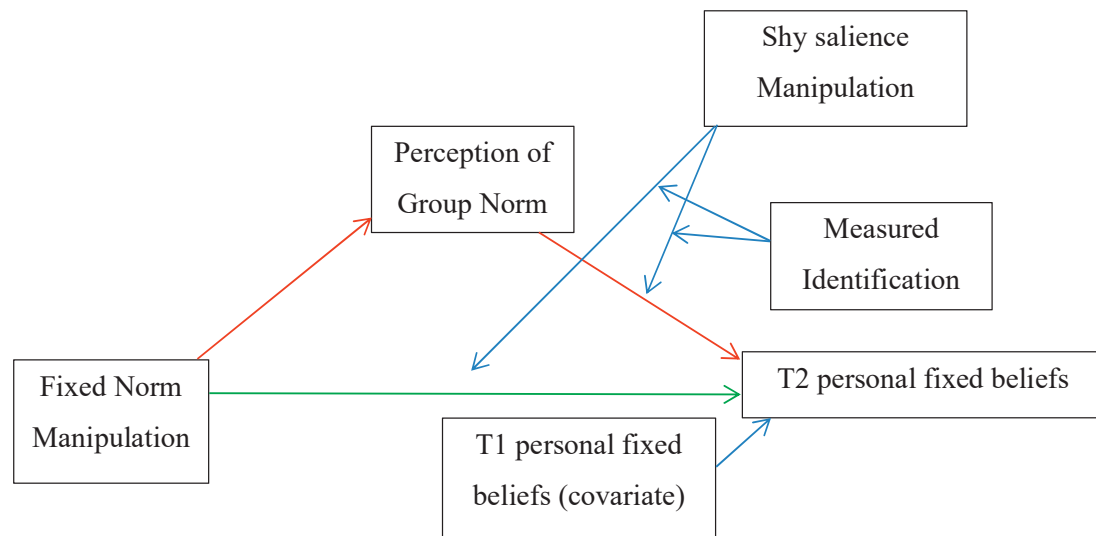


Figure 10: Conditional mediation model comparing social influence and social differentiation explanations.

The model compares the: conditional indirect effect (indicated in red and consistent with social influence) whereby the norm manipulation predicts perception of the group norm, which goes on to predict T2 fixed beliefs, conditional on identification and shy salience; and a conditional direct effect (indicated in green and consistent with social differentiation) whereby the fixed norm manipulation, over and above the mediated effect, predicts T2 fixed beliefs, conditional on identification and shy salience. (See Model 19, Hayes, 2012)

Before running this analysis, it was important to test a key assumption of this proposed model according to which participants across all conditions rated the perception of the group norm in a similar way. To test this assumption, participants' ratings on the continuous word-cloud manipulation check (i.e. the "Perception of the group norm" mediator in Figure 10) served as the dependent variable. The perception of the group norm was regressed on group salience, the group norm, measured identification, and again included all two and three way interactions. This revealed, as expected, a significant main effect of the norm manipulation ($p < .01$), and no other main or interactive effects (all $ps > .05$). This confirmed the assumption that the fixed and the malleable word clouds lead all participants to

perceive the group norm to be fixed and malleable respectively, and to a similar degree, irrespective of participants' level of identification or which salience condition they were in.

Next, to test the main conditional mediation model (shown in Figure 10), bias corrected bootstrapping was used to generate 95% confidence intervals for the direct and indirect effects. If confidence intervals do not include zero, then this indicates a statistically significant effect. Results based on 10 000 bootstrap samples showed a significant indirect effect. For high identifiers (1 SD above the mean), under group salience conditions, the confidence interval for the indirect effect did not include zero (IE = .23, lower 95% CI = .04, upper 95% CI = .44), but for low identifiers, under low salience conditions, the indirect effect was not significant because it did include zero (IE = -.26, lower 95% CI=-4.86, upper 95% CI=2.73). This finding indicated that the indirect effect, via perceptions of the group norm, was only significant for high identifiers under high group salience conditions. This suggests this subgroup was influenced by what they perceived the group norm to be, consistent with social influence.

The direct effects of the norm manipulation, over and above the indirect effect, were then assessed. This revealed that only for low identifiers, under low group salience conditions, was there a significant direct effect (DE = 6.83, lower 95% CI=-2.53, upper 95% CI=16.20). This finding suggested that low identifiers, under low group salience conditions, were differentiating themselves from the group norm.

Study 2 discussion

Overall, Study 2 did not provide support for Empirical Proposition 1 (and the individual mobility thesis) or Empirical Proposition 2 (and the collective mobility thesis) when experimentally manipulating identity salience and the group norm. However, follow-up analysis revealed that group salience and the group norm were interactively affecting participants' personal fixed beliefs but in different ways, depending on participants' level of

identification with the shy group. Partially consistent with Empirical Proposition 2 and Corollary 2 (and the broader collective mobility thesis) there was evidence that high shy identifiers, under high group salience conditions, perceived their condition as more malleable when the group expressed a more malleable norm compared to when the group displayed a more fixed norm. However, contrary to Empirical Proposition 1, and the individual mobility thesis, there was no main effect suggesting that the individual identity salience condition (alone or in interaction with identification) leads individuals to perceive their shyness to be more malleable. Unexpectedly, when low identifiers, under low salience conditions, were exposed to a malleable group norm they appeared to strongly endorse personal malleability beliefs.

There are at least two possible interpretations of this latter effect. First, it could be that under these conditions, individuals were socially influenced by the malleable word cloud. Such an interpretation conflicts with extensive social identity research showing that social influence is an outcome of a shared identity (Turner, 1991; 2010). An alternative explanation was that low identifiers were differentiating themselves from the group norm.

This latter interpretation was supported by the follow-up conditional mediational analysis. Strikingly, this analysis revealed that low identifiers, under low group salience conditions, understood what the group norm was (and in a similar way to other participants) but this subgroup did *not* simply conform to this norm, since there was no significant effect of the perception of the group norm on their individual beliefs. Rather, the malleable word cloud appeared to lead low identifiers, under low salience conditions, to rate their own condition as more malleable, over and above what they perceived the group norm to be.

By way of contrast, under high salience conditions, high identifiers' perceptions about what the group believed went on to predict their personal beliefs about the relative malleability of their personal shyness. In other words, high identifiers, under high group

salience conditions, appeared to conform to the group norm whereas low identifiers, under low group salience conditions, appeared to differentiate themselves from the group norm.

Some limitations to this analysis should be noted. There was no evidence that the salience manipulation alone was sufficient to induce members to be influenced by ingroup members. Perhaps, in retrospect, this should come as no surprise. SCT posits that the salience of a particular identity is an outcome of both the fit of a self-category in a situation as well as the perceiver's readiness to use a particular self-category (which includes a person's sense of long-term identification). In Experiment 2, then, it is likely that the effect of long-term identification with the group played a far more substantial role in making a shy identity salient. Nevertheless, it is notable that the effect of the norm manipulation was observed only when high identifiers were exposed to the group salience manipulation. Overall, then, it seems that high group salience plus low identification and low group salience plus high identification had a net, neutral effect on group level self-categorisation and thus attenuated any subsequent social influence or differentiation effects for these subgroups.

These findings also highlight the methodological difficulties in experimentally manipulating identity salience and thereby providing causal evidence that the proposed Disorder Identity processes affect people's personal malleability beliefs. This leaves open the question as to whether there is not some unobserved correlate of social identification that is driving the observed pattern of results. Of particular relevance to the current enquiry, is the question of whether, contrary to the collective mobility thesis, it is not identification with fellow disorder sufferers *per se* that matters, but rather identification with any fellow ingroup members that can underpin the influence process. This important issue was explored in Experiment 3.

Study 3 introduction

While the results of Study 1 and Study 2 both provided some evidence for the collective mobility thesis, these results do not speak to the claim that fellow sufferers are uniquely informative about the malleability of a condition. A key element of the collective mobility thesis is that fellow sufferers provide validated and theoretically relevant experience. By contrast, fellow ingroup members from other groups to which an individual belongs, may be similar in many ways, and may well have beliefs about the malleability of a condition, but this is less likely to inform an individual's personal beliefs about the malleability of a condition (i.e. Empirical Proposition 3). To test these ideas, in Study 3 I examined to what extent shy individuals would be socially influenced when presented with a normative understanding that shyness is malleable from either irrelevant fellow ingroup members or relevant ingroup members.

Study 3 method

Participants and design

Seventy-three eligible study participants (median age = 22 years; 43 females) who reported experiencing shyness symptoms completed the study as part of the coursework component of a third year psychology laboratory class at the Australian National University (ANU). The study design did not involve manipulating the group norm, but rather manipulated the relevance of the ingroup expressing the view that shyness was malleable. The ingroup was presented as either fellow shy group members (relevant ingroup) or fellow ANU students (irrelevant ingroup). Participants' levels of identification with the group of shy people and with the group of fellow ANU students served as measured, predictor variables. As in the previous studies, the dependent variable was participants' individual fixed beliefs at Time 2, measured after the group relevance manipulation, controlling for baseline (Time 1) individual fixed beliefs.

Materials and procedure

Participants were first asked “how shy are you”, and asked to respond using the same 1 (“not at all shy”) to 5 (“extremely shy”) Likert scale used in the earlier studies. Participants who reported levels of shyness below the midpoint of the scale ($N = 43$) completed the study but their results were not included in the final analysis.

Eligible study participants were then asked a set of more in-depth questions assessing their shyness symptoms, using the 13 item Revised Cheek and Buss Shyness Scale (RCBS), used in the previous studies. Participants then rated their personal beliefs about the malleability of shyness using the same five item measure used in the previous studies. Participants were then asked to rate their identification with the group of shy people using the same, seven-item measure used in the previous studies. However, participants were also asked to rate their level of identification with fellow ANU students, using the same social identification questions. The order of these two blocks of questions, assessing ANU identification and shy group identification, were randomly counterbalanced.

After completing these introductory questions, participants were introduced to the main ingroup relevance manipulation. This manipulation aimed to alter whether participants were exposed either to ANU students’ beliefs about shyness (irrelevant ingroup) or fellow shy people’s beliefs about shyness (relevant ingroup). In both conditions, participants were presented with the same “word cloud” used within the malleable group norm condition used in previous studies. However, in the ANU group condition, participants were told: “We’ve used special software to analyse selected keywords from previous research asking Australian National University students to discuss what they think about the nature of shyness.” In the shy group condition, participants were presented with the same information except they were told the word cloud was from shy people discussing shyness, rather than ANU students discussing shyness.

Participants were then asked a manipulation check asking whether “The word cloud reflects the opinions of...” with the response options of either “a group of shy people” or “a group of ANU students”. To further reinforce the manipulation, participants were asked to indicate their level of agreement with the statement “I believe the word cloud shows how (shy people / ANU students) believe that: ” and then, in both conditions, presented with four statements from Beer’s (2002) implicit theory of shyness measure. The four statements included two malleable items (e.g. “Shy people can change aspects of their shyness if they want”) counterbalanced with two fixed items (e.g. “Shy people have a certain level of shyness and it’s something they can’t do much about”). To reinforce the identity of the discussants, the words “ANU students” and “shy people” (in the initial stem only) were bold and underlined.

Finally, participants were again asked the five items assessing individual malleability beliefs at Time 2 used in the previous studies. Participants were then asked to indicate their age and gender and were debriefed about the nature of shyness.

Study 3 Results

Manipulation checks and preliminary analyses

Analysis of the manipulation check question revealed that 13 of the 36 participants in the ANU identity condition and 6 of the 37 participants in the shy identity condition incorrectly indicated whose opinions were reflected in the word cloud. However, subsequent analyses, testing the main study hypotheses, were substantively the same whether those who failed the manipulation check were included in the analysis or not. This may have been because after completing the manipulation check, but before the main dependent variable, the study procedure included questions further emphasising the identity of the discussants (either ANU students or shy people). This suggested that participants’ understanding of the main manipulation may have consolidated after the main manipulation check. For this reason, the

main analysis will be reported for the full sample ($N = 73$), including participants who failed the manipulation check. For the main study hypotheses, results will also be reported separately for the subsample who answered the manipulation check correctly ($N = 54$).

The 13 items from the RCBS (Cronbach's $\alpha = .86$), the five items from the personal fixed beliefs about shyness at Time 1 (Cronbach's $\alpha = .80$) and Time 2 (Cronbach's $\alpha = .77$), and seven social identification items assessing identification as shy (Cronbach's $\alpha = .89$) and identification as an ANU student (Cronbach's $\alpha = .87$) all formed reliable scales. To assess unidimensionality, the items within each scale were subject to a principal components analysis. The items assessing Time 1 and Time 2 fixed beliefs, and shy identification items each loaded on a single component with an eigenvalue greater than 1. The items assessing ANU identification loaded on a single component with an eigenvalue of 3.90 explaining 55.70% of the variance and then a second component with an eigenvalue of 1.25 explaining an additional 17.82% of variance. However, visual inspection of the associated scree plot showed that the components levelled off after the first component, indicating that the scale was likely to be unidimensional. Similarly, the RCBS loaded on a single component with an eigenvalue of 5.14, explaining 39.63% of variance and there were three additional components with eigenvalues less than 1.44 but greater than 1, explaining a further 27.83% of variance. Again, visual inspection of the associated scree plot suggested the components levelled off after the first component, indicating that the items were unidimensional. As in the earlier studies, the RCBS score was calculated by reverse scoring negatively worded items and then summing across all items. The identification and fixed belief measures were calculated by taking the mean of the item scores for each measure after reverse-scoring malleable belief items.

The means, standard deviations and bivariate correlations between the main study variables are shown in Table 10. As in the earlier study, the mean RCBS score was more

than one standard deviation higher than the mean reported in norms developed by Hopko et al. (2005), indicating that the screening question successfully selected participants with above average levels of shyness symptoms.

Table 10: *Study 3, means, standard deviations, and bivariate correlations between main study variables.*

Measure	M	SD	1.	2.	3.	4.	5.
1. Revised Cheek and Buss Shyness Scale (RCBS)	40.29	7.13	--	0.22	.26*	-.035	.24*
2. Personal Fixed Belief (T1)	3.68	0.66		--	.30**	0.09	.75**
3. Shy identification	3.69	1.06			--	.37**	.45**
4. ANU identification	4.49	0.87				--	0.11
5. Personal Fixed Belief (T2)	3.54	0.62					--

** correlation is significant at the .01 level (2-tailed), * correlation is significant at the .05 level (2-tailed)

Study 3 main analysis

To test Hypothesis 3, that only those who identified as shy would be influenced by fellow shy people about the malleability of shyness, a 2 (group relevance: ANU, shy people) x 2 (identification as shy: higher, lower) x 2 (identification as an ANU student: higher, lower) analysis of covariance was run. Group relevance was a manipulated independent variable, identification as shy and identification as an ANU student were both continuous (mean-centred) independent variables. As in the earlier studies, Time 1 personal fixed beliefs was included as a continuous covariate, and Time 2 personal fixed beliefs was the dependent

variable. The hypothesised two-way interaction between group relevance and shy group identification was significant ($F(1,72) = 4.93, p = .03$) and there was also a main effect for the shy group identification ($F(1,72) = 7.22, p < .01$), and the covariate ($F(1,72) = 81.56, p < .001$). The interaction effect was probed by analysing the main effect at one standard deviation above (high identifiers) and below (low identifiers) the mean level of social identification as shy, while holding ANU identification and Time 1 personal fixed beliefs constant at their respective mean values. None of the simple effects were significant (all $ps < .05$). However, inspection of means, as shown in Figure 9, revealed that high shy identifiers tended to perceive their own shyness to be less fixed when fellow, shy ingroup members were discussing the malleability of shyness compared to when fellow ANU students were discussing the malleability of shyness. By contrast, the opposite pattern was observed among those with lower shy identification. Together these findings indicated that for high shy identifiers, fellow shy ingroup members discussing the malleability of shyness influenced personal malleability beliefs, but fellow ANU ingroup members discussing the malleability of shyness did not similarly influence personal malleability beliefs. Conversely, low identifiers' personal malleability beliefs were lower overall but tended to be more malleable when ANU students were discussing the malleability of shyness rather than when shy people were discussing the malleability of shyness.

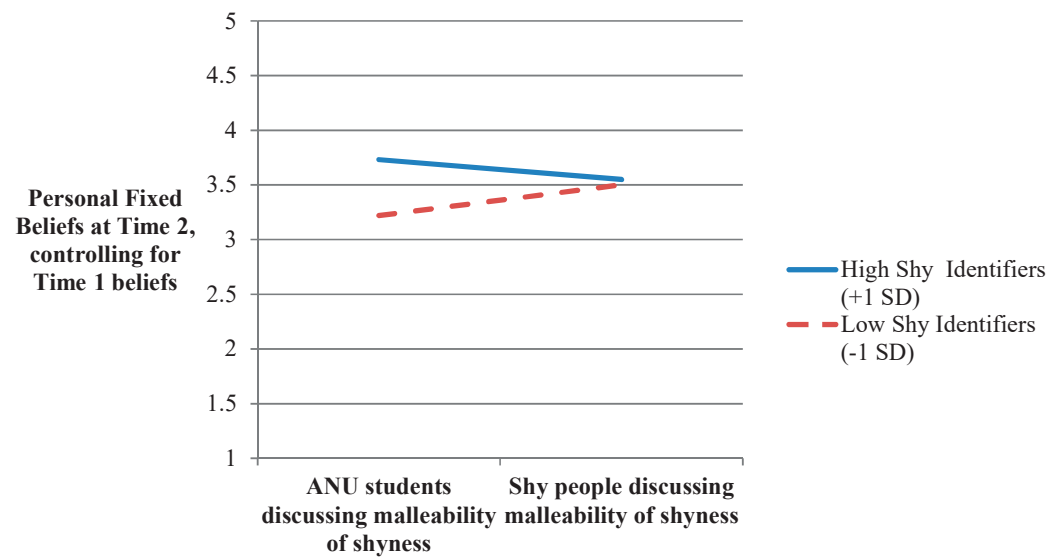


Figure 9: Significant two-way interaction between shy group identification and manipulated malleability norm on personal fixed beliefs at Time 2 (controlling for Time 1 fixed beliefs). Only highly identified shy group members were influenced by the malleable norm.

Finally, the main analysis was repeated using the subset of participants who correctly answered the manipulation check. This again revealed a significant two-way interaction between group relevance and shy group identification ($F(1,53) = 4.88, p = .03$) and the main effect for shy group identification ($F(1,53) = 4.43, p = .04$), and the covariate ($F(1,53) = 40.63, p < .001$).

Study 3 discussion

Study 3 provided important evidence consistent with Hypothesis 3, and the broader collective mobility thesis, that fellow Disorder Identity group members can be an important source of social influence with respect to disorder malleability beliefs because they provide validated information about the nature of shyness. Conversely, people with low shy group identification, if anything, displayed the opposite pattern of results. In other words, it was not the case that “people like me”, in any respect, provided a basis for informing my beliefs about

the malleability of shyness. What matters, is that people who are “like me” in a theoretically relevant way inform my beliefs about my capacity to change.

Although the purpose of the present study was not to test Corollary 1 and the individual mobility thesis, it was notable again that identification as shy was positively correlated with more fixed personal beliefs about shyness. Interestingly, in the final model there was also a significant main effect for shy identification on fixed beliefs at Time 2, after controlling for Time 1 beliefs. It was unclear what might explain these findings, other than the fact that the enduring focus on shyness increased the overall salience of the shyness self-category over the course of the study for those who identified as shy to begin with.

There were a number of limitations to the study that should be noted. Firstly, the sample size, after screening for those who experienced significant shyness symptoms, was relatively small given the design of the study, thus increasing the risk of Type 1 error. It was also unfortunate that such a large number of participants failed the main manipulation check, even if the results were substantively the same whether or not these participants were included in the analysis. However the basic pattern of results is unlikely to reflect a chance association given that the results were very consistent with the findings from Study 1 and Study 2.

General discussion

Together, the findings from Studies 1, 2 and 3 start to address some important questions about the Disorder Identity analysis in general and the collective mobility and individual mobility theses in particular. All three studies provided consistent evidence suggesting that those who identify as shy, when presented with evidence that fellow shyness sufferers endorse a malleable norm, subsequently perceived that their own, personal shyness was more malleable. These findings align with the overarching collective mobility thesis, and were specifically consistent with Corollary 2 and Empirical Proposition 3. The findings

from Study 2 provided only partial support for Empirical Proposition 2 by showing that by experimentally manipulating shy salience, participants were influenced by the shy group malleability norm, but only for high identifiers.

However, there was also evidence consistent with Corollary 1 and the individual mobility thesis: identification with the shy group was consistently significantly associated with more fixed Time 1 beliefs about shyness in Study 1 ($r = .44$), Study 2 ($r = .31$), and Study 3 ($r = .30$). Conversely, of course, these findings suggest that low identification as shy is associated with more malleable beliefs. Interestingly, contrary to Empirical Proposition 1, there was no evidence that experimentally manipulating the salience of people's personal identity, compared to making salient their shy identity, caused participants to hold more malleable beliefs. Unexpectedly, however, Study 2 revealed that low identifiers, under low salience conditions, appeared to hold particularly malleable beliefs when presented with a malleable group norm. The results from the conditional mediation analysis suggested that this was not because participants conformed to what they perceived the group to be saying (i.e. social influence), but rather because they appeared to be positively differentiating themselves from the shy group.

However, Study 2 does not speak to the question of why there was no evidence for a main effect of the personal identity salience manipulation on personal malleability beliefs (i.e. Empirical Proposition 1), despite correlational evidence for Corollary 1. It could be because there is no causal relationship between identity salience and personal malleability beliefs, or it could be because people's prior identification as shy overwhelmingly drives the self-categorization process compared to the rather brief salience induction used in Study 2. The latter interpretation seems more plausible given the findings of Ellemers et al. (1997), showing that by manipulating group identification, using a more elaborate "bogus pipeline" procedure, low identifiers tended to prefer an individual upward mobility strategy.

Conceivably, then, the salience manipulation was simply insufficiently robust to be able to drive the hypothesised effects. Consequently, the findings reported here, while suggestive, do not provide causal evidence to support the individual mobility thesis or the collective mobility thesis.

However, Study 3 did help rule out an alternative account of the patterns observed in Study 1 and Study 2. Study 3 showed that individuals who identified as shy tended to believe their own shyness was relatively more malleable when fellow shy sufferers displayed malleable beliefs compared to when fellow university students expressed the same views. By contrast, low shy identifiers, if anything displayed the opposite pattern of results. This is an important point to consider in the context of the broader Disorder Identity analysis because it is consistent with the claim that it is not merely group processes in general that shape health and illness beliefs (Jetten et al., 2011), but that Disorder Identities specifically play an especially important role. The findings underscore the point that fellow sufferers, who are like me in a theoretically relevant way, provide an epistemic resource for informing me about who I am and the constraints of my condition. By contrast, my individual experience of setbacks and failure in trying to manage or perhaps overcome my shyness provides a comparatively noisy dataset on which to inform my beliefs about my prospects for change.

The point further echoes a recurring theme throughout this thesis: the claim that categorization is fundamentally a meaning-making process (McGarty, 1999) that serves to predict, explain and control. But, as noted in Chapter 1 and Chapter 4, the categorization process is not constrained by mere similarity alone (Murphy & Medin, 1985)—it is also determined by, among other things, background theories and knowledge (Ahn et al., 1995, Turner & Brown, 2002). This point goes some way to explaining why it is that social modelling of achievement by “similar others” (in terms of age and gender, for example) is

only inconsistently related to self-efficacy beliefs in a particular domain (Bandura, 1997, Usher & Pajares, 2008).

More broadly, the present findings point to the diversity of ways in which people's beliefs about the malleability of a condition can change as a function of self-categorization processes. It should be noted that the studies did not seek to test or compare the role of professional expertise in shaping malleability beliefs, and nor do the findings discount or conflict with the claim that professional expertise can be profoundly influential in shaping malleability. From the perspective of SCT, however, such influence reflects the fact that professional experts are often considered to be theoretically "relevant" sources of knowledge for informing me about my condition because I identify with a wider societal group that values modern medical science and views medical experts as legitimate representatives of "us" in these domains. However, as noted by Haslam (2005), lay theories about mental disorder are not merely attempts to approximate expert empirical understandings, in part because diverse social-psychological processes are involved in governing these lay understandings.

To summarise this point, extant social-psychological theorising suggests that implicit theories about the malleability of a condition are primarily shaped by: individual differences (Chiu et al., 1997), early learning experiences (Dweck & Elliott-Moskwa, 2010), informational influence from experts (Blackwell et al., 2007), or personal identity salience / low group identification (Ellemers et al., 1997). The current findings provided support for the latter, individual mobility thesis. But the findings also suggested that there are at least two further ways that an individual's personal malleability beliefs can change: by differentiating the self from the group in circumstances where low group identifiers' personal identity is salient, or by identifying with fellow group members who perceive that a group-defining dimension is malleable (i.e. social influence).

It is the evidence for this latter point that provides an important first step in assessing the basic claim being explored in the empirical chapters of this thesis that disorder identification need not be a barrier to change. However, this evidence consistent with the collective mobility thesis must also be tempered with the acknowledgement that, consistent with previous research (Yzerbyt et al., 2000), identifying with the group is also strongly and consistently correlated overall with fixed personal beliefs. Finally, it is also important to be able to demonstrate not just that group processes shape personal beliefs about malleability but that these beliefs go on to predict variation in symptomatology and behaviour—a point taken up in Chapters 7 and 8.

Chapter 7: Study 4 – The Interaction Between a Malleable Group Norm and Disorder Identification Predicts Reductions in Symptomatology Over Time.

Studies 1-2 demonstrated that people who experience symptoms of shyness can be influenced by other shy people about the malleability of their own shyness, but only if they psychologically identify as shy. While this begins to provide an important part of the evidence base for understanding how Disorder Identities might change people's beliefs about who they are and who they can become, an important next step is to demonstrate that these malleability beliefs have behavioural and symptomatic consequences beyond the laboratory. Indeed this is a central claim implied by the Disorder Identity analysis: collective understandings of a condition, in terms of a shared Disorder Identity, are what systematically changes behaviour and symptom perception. Specifically, in the context of the collective mobility thesis, the claim is that to the extent people identify with a disorder category, and the group holds a norm of growth and change, then it is predicted that this identity will serve as a negatively homeodynamic and positively heterodynamic mechanism. The purpose of the present chapter, therefore, is to test Empirical Proposition 4: that to the extent individuals identify with a disorder category, and there is a malleable group norm, then it is predicted that those individuals will display reductions in symptomatology over time.

Research is increasingly demonstrating the important role that implicit theories play in influencing the aetiology and maintenance of symptoms. People's beliefs about the malleability of their personal attributes change their motivations to engage in therapy (Dozois & Westra, 2005; Petrie, Weinman, Sharpe, & Buckley, 1996), their self-regulation (Burnette, O'Boyle, Vaneppe, Pollack, & Finkel, 2012), their illness attributions and their coping behaviours (Dweck & Elliott-Moskwa, 2010). Over and above the impact of prior symptom severity, individuals who hold a more fixed view tend to experience worse clinical outcomes across a range of clinical and sub-clinical conditions, including depression (Dweck & Elliott-

Moskwa, 2010), weight loss (Burnette et al., 2012), and shyness (Beer, 2002; Valentiner, Jencius, Jarek, Gier-Lonsway, & McGrath, 2013; Valentiner, Mounts, Durik, & Gier-Lonsway, 2011). Beer, for instance, demonstrated in a series of three experiments that individuals who experience shyness, and held a relatively malleable theory of shyness, were more inclined to prefer a difficult learning goal to help overcome their shyness, showed a greater willingness to approach social settings, and experienced fewer negative emotional consequences following social interaction.

These findings illustrate the clinical promise of interventions seeking to foster malleable self-theories, particularly in the context of disorders where motivational processes play a central causal role in maintaining the disorder. However, the relationship between self-theories and the temporal dynamics of symptom change are somewhat unclear. The measures used in implicit theory research provide little or no context about the durability and time-scales over which change might be expected to occur. Two of the items, for instance, ask respondents to rate their agreement with the statements: “How (shy) I am is something that changes through my life” and “My (shyness) is not fixed but changes over time”. But these items do not specify the *nature* of these changes with respect to their magnitude, rapidity, or stability, for instance. The implication of this is simply that group norms about the malleability of a condition, such as those based around Dweck’s (1997) implicit theory measures, are likely to require more contextual information about how and when changes occur, in order to make more specific behavioural predictions.

One such context, where these expectations are likely to be more explicit, is the transition to a novel social environment. Starting university, for example, is a situation in which there is a shared social expectation that a certain amount of social anxiety, or shyness, is quite normal, and perhaps even socially advantageous (Valentiner et al., 2011). Individuals experiencing shyness during this transition can, therefore, plausibly make a mix of internal

and external attributions for explaining their feelings of shyness. Over the course of a university year, there may be a reasonable expectation that the experience of shyness is socially normative but attenuates, as students become more familiar with the environment and build new friendships and confidence in their social skills. But for those individuals who perceive themselves to be shy to begin with, these external attributions may be tempered with beliefs that their feelings may also reflect something dispositional *about themselves*. These competing understandings are likely to play out in daily opportunities for social approach and avoidance. Should I speak up in a tutorial or stay quiet? Should I stay in my dorm room or go to the common room? Should I talk to somebody in the queue at the refectory or not?

As shown in Studies 1-3, the beliefs, and implicitly the experience, of relevantly similar others, is an important determinant of these decisions. In the context of transitions to university those who are gregarious from day one are clearly not “like me”. But what of those who seem to be *changing*—those who start out shy like me, but appear to be becoming gradually more sociable? Quite divergent attributional inferences can be drawn about why these individuals are changing. Perhaps they were never *really* shy to begin with. Perhaps their shyness can be overcome but mine cannot. Or perhaps their experience informs my lay beliefs about whether I too can change.

The thrust of the collective mobility argument developed so far is that, following SCT (Turner et al., 1994), to the extent that an individual self-categorizes as part of the social group of fellow shyness sufferers, the self becomes depersonalised, and fellow group members become cognitively interchangeable with the personal self. But the “fit” of this self-category depends on both normative and comparative elements: putative ingroup members, who in a particular situation display relatively low comparative fit (for instance, because their behaviour is less socially anxious than mine), may nevertheless be perceived to be part of the group to the extent that their behaviour is *normatively* fitting (because *we*

believe that incremental change is possible). In turn, these people's behaviours become a source of valid information, and perhaps inspiration, that I too might change. Conversely, to the extent individuals perceive that the shy group has a more fixed view about shyness, the putative group members who appear to overcome their shyness, and so display both poor normative and comparative fit, are likely to be perceived as "other". Consequently, their experience is not a valid source of information about who I could become.

These theoretical claims suggest that normative and comparative fit play an ongoing role in interactively determining each other over time. But empirically, it is true that there is little if any research, and certainly not in the clinical domain, examining how identity and normative fit interact over time. Social identity researchers have run longitudinal studies demonstrating the health benefits of socially identifying with groups that are unrelated to Disorder Identities, in coping with stress (Haslam, Jetten, & Waghorn, 2009), and in improving physical health (Khan et al., 2014). But in both cases, these studies assume that the underlying mechanism reflects how group memberships, regardless of their defining attributes, provide a generic source of socio-emotional support and thereby buffer stress by moderating primary and secondary appraisals (Haslam, Jetten, O'Brien, & Jacobs, 2004). Neither of these findings speaks to the specific processes involved for groups that are defined by a shared experience of symptoms.

The argument that is of interest here is therefore importantly different. The claim is that Disorder Identities provide specific and relevant information about how my symptom experience is likely to change, and that this in turn will alter how symptoms are experienced. Specifically, in the context of shyness, Empirical Proposition 4 was the claim that to the extent that individuals identify as shy, and there is a malleable group norm, then those individuals will display reductions in shyness symptoms over time. The present study sought

to test this by examining changes in social anxiety symptoms over a six month period within a group of first-year university students who self-reported experiencing shyness.

Study 4 method

Participants and procedure

Participants were recruited using online and campus message boards at the Australian National University (ANU). Participants were invited to complete an online questionnaire investigating “shyness and anxiety in the transition to university”, in exchange for either \$5 or 30 minutes of course credit (for first-year psychology students). Recruitment advertisements made it clear that eligible participants needed to be in their first-year of study at ANU, have personal experience with shyness since starting at ANU, as well as a willingness to complete a follow-up survey in six months’ time. One hundred and fifty-five eligible participants completed the study at Wave 1 (median age = 18 years; 105 females). Participants completed the first wave of the study an average of 42 days after the start of first semester of first-year.

Time 1 participants were then contacted via email an average of 204 days later and invited to complete the final wave of the survey. Four follow-up emails were sent to participants encouraging them to complete the second survey and reminded that they would again be offered \$5 for completing the study. Unfortunately, only 37 participants (24% of Time 1 respondents) completed the Time 2 survey (median age = 18 years, 27 females).

Measures

Participants answered all questions through a web-based survey software link. Participants first answered a series of questions assessing their eligibility for the study. Participants were asked to confirm that they were a first-year university student and then responded to the question “how shy are you?” on a 1 (“not at all shy”) to 5 (“extremely shy”) Likert scale. The online survey software was programmed to screen for participants who

were first-year university students and who rated their shyness at 3 or above on this item. Participants who reported levels of shyness below the midpoint of the scale ($N = 67$) were thanked for their time and did not answer any further questions.

Eligible study participants then completed the same 13-item Revised Cheek and Buss Shyness Scale (RCBS) used in Studies 1-3. Participants then completed a 20-item measure of social performance anxiety, the Social Interaction Anxiety Scale (SIAS), successfully used by Valentiner et al. (2011) in their study assessing shyness in the transition to university. This scale included items such as “I feel tense if I am alone with just one person”, to which participants responded using a 1 (“Not at all true or characteristic of me”) to 5 (“Extremely characteristic or true of me”) Likert scale. Participants then rated their level of agreement with a 5-item measure assessing personal fixed beliefs about shyness (Beer, 2002), that was the same as used in Study 1-3. Participants rated their level of agreement to this, and the subsequent measures, using a 1 (“Strongly Disagree”) to 7 (“Strongly Agree”) Likert scale.

Participants rated their level of identification with the group of “shy first-year ANU students” by responding to all but one of the items used in Studies 1-3 to measure social identification (i.e. the social identity measure based on Leach et al. (2008), including items such as “I am similar to the average person in this group (of shy first-year ANU students)”). The single item that was not included was the item “Being a shy first-year ANU student is an important part of how I see myself”. This decision was taken in part because the wording “shy first-year ANU student”, unlike the categories “shy people” or “ANU student”, was likely to be a comprehensible but seldom used category label. It was decided to not include this item in the questionnaire because seeking agreement that this unusual category was “an important part of how I see myself” seemed awkward and potentially likely to undermine the credibility of the study.

Participants then answered six items measuring the perceived group norm regarding the malleability of shyness, again using modified items from Beer (2002) (e.g. “Shy first-year ANU students believe they have a certain level of shyness and it is something they can’t do much about”).

Participants then rated their social identification with all students at ANU, using the same items used to measure identification with the group of shy first-year ANU students. This measure was included because the main identification measure “shy first-year ANU students” did not merely tap into the category “shy people”, as in earlier studies, but rather the subgroup of “shy, first-year ANU students” and therefore arguably confounds the interpretation of the hypothesised norm by identification interaction. Also, previous research by Valentiner et al. (2011) suggested that the related construct of college belongingness had a significant effect on changes in shyness over the course of the transition to university. By measuring the closely related construct of ANU identification it would be possible to test this alternative account. Finally, participants answered questions assessing their age and sex, before being reminded about the follow-up survey in six months’ time and being asked to provide their email address. In the Time 2, follow-up survey, participants answered the same measures, excluding the screening and demographic questions.

Study 4 results

Preliminary analyses

The 13 items from the RCBS at Time 1 (Cronbach’s $\alpha = .86$), SIAS at Time 1 (Cronbach’s $\alpha = .86$) and Time 2 (Cronbach’s $\alpha = .87$), the five items from the personal fixed beliefs about shyness at Time 1 (Cronbach’s $\alpha = .67$), six shy group identification items at Time 1 (Cronbach’s $\alpha = .90$), ANU group identification items at Time 1 (Cronbach’s $\alpha = .79$), and six items measuring the perceived group norm that shyness is fixed (Cronbach’s $\alpha = .88$), all formed reliable scales.

To assess the unidimensionality of the focal study variables, a series of principle components analyses were run on the items measuring SIAS at Time 1 and Time 2, shy identification and the perceived group norm that shyness is fixed. This revealed that the items measuring the group norm of shyness and shy group identification each loaded on a single eigenvalue greater than 1. SIAS at Time 1 and Time 2 each loaded on four components with eigenvalues greater than 1. However, visual inspection of the associated scree plots revealed a single, dominant component, followed by a levelling off of the subsequent components, indicating that these measures were unidimensional.

The RCBS score was then calculated by summing individual items within this scale (after reverse scoring negatively worded items). For the other study variables the scale scores were calculated by taking the mean of the individual items (again, after reverse scoring negatively worded items). Table 11 shows the means and standard deviations of, and correlations between, the main study variables.

Table 11: Study 4, means, standard deviations and bivariate correlations between main study variables.

Measure	M	SD	1.	2.	3.	4.	5.	6.	7.
1. Revised Cheek and Bus Shyness Scale (RCBS, T1)	42.35	7.57	--	.47**	-.03	.37*	-.40*	.71**	.52**
2. Personal Fixed Belief (T1)	3.53	.85		--	-.07	.36*	-.28*	.34*	.38*
3. Shy First Year ANU Student Group Identification (T1)	4.12	1.21			--	-.28	.67**	.04	-.17
4. Fixed Group Norm (T1)	3.19	1.06				--	-.26	.10	.19
5. ANU Identification (T1)	4.30	1.00					--	-.32	-.40*
6. Social Interaction Anxiety Scale (SIAS, T1)	2.38	.69						--	.77**
7. SIAS (T2)	2.13	.70							--

* Correlation significant at p<.05 (two-tailed). ** Correlation significant at p<.01 (two-tailed).

The mean RCBS score ($M = 42.35$, $SD = 7.57$) was more than one standard deviation higher than the mean reported in norms developed by Hopko et al. (2005), indicating that the screening question successfully selected participants with above average levels of shyness.

Noting the high level of attrition, an analysis was run to assess if there were any significant differences between those participants who did and did not complete the Time 2 survey, with respect to the main baseline measures shown in Table 1. Between-subjects t -tests revealed only one statistically significant difference: those who completed the Time 2 survey had significantly higher shyness scores ($M = 42.35$) compared to non-completers ($M = 39.23$, $t(156) = 2.45$, $p = .015$). This suggested that while the high level of attrition was unfortunate, there was no evidence that differential attrition might systematically confound the interpretation of any findings (as might arguably be the case if non-completers had significantly more fixed beliefs about shyness, or had higher shyness levels). Indeed, if anything, the final sample was *more* shy than non-completers.

Main analysis

Before assessing the main hypothesis, it was important to determine whether, among those who completed the Time 2 measures, there were any baseline differences in SIAS levels as a function of the main predictor variables. To assess this, Time 1 SIAS was regressed against identification with shy first-year ANU students, the perceived group malleability norm and the (mean-centred) interaction. This revealed no significant effects (all $ps > .05$), indicating that there was no evidence that SIAS varied systematically as a function of the primary independent variables.

To test the main hypothesis, a repeated measures ANOVA was run with SIAS at Time 1 and Time 2 as the repeated measure, and Time 1 identification with the group of shy ANU students, the Time 1 group malleability norm, and the interaction term, as continuous predictor variables. The reason for using a repeated measures ANOVA, rather than

ANCOVA as in the earlier studies, was that such an approach more accurately reflects the hypothesised effect of intra-individual change in symptomatology. Such an approach can also provide greater statistical power. For interest, the analyses were re-run using ANCOVA, and although the pattern of means was in the same direction as reported below, the hypothesised interaction effect was not statistically significant ($p = .14$)

The repeated measures ANOVA revealed a significant three-way interaction between the repeated measure, group norm and group identification ($F(1,33) = 4.32, p = .05$), and a main effect for identification ($F(1,33) = 4.87, p = .03$). To probe the interaction, a difference score was calculated by subtracting the Time 1 SIAS score from the Time 2 SIAS score and regressing this against the main and interactive effects of the group norm and identification. The results of this analysis are arithmetically identical to the repeated measures ANOVA (Weisberg, 1980) but facilitate simple slopes analysis (Preacher, Curran, & Bauer, 2007) of the differences at one standard deviation above and below the mean social identification and norm scores. This simple slopes analysis revealed that for those who scored one standard deviation below the mean on the fixed norm measure, there was a significant relationship between shy identification and the reduction in SIAS between Time 2 and Time 1 ($t(33) = 2.80, p < .01$). However, for those who scored one standard deviation above the mean on the fixed norm manipulation there was no significant relationship ($t(33) = .47, ns$). Figure 11 illustrates how those who highly identified with the group of shy ANU first-year students, and perceived that the group norm was malleable (i.e. -1 SD below the mean on the fixed norm measure) experienced a substantial reduction in social interaction anxiety between Time 1 and Time 2. This confirmed the hypothesis that high shy group identifiers, who perceived that the group norm was relatively malleable, experienced the largest decline in their personal SIAS levels across the two time periods.

One possible confounding element in this analysis was that participants' perceptions of the group norm simply reflected their personal malleability beliefs, rather than their beliefs about the group's beliefs. To test this possibility, the analysis was repeated but adding Time 1 personal fixed beliefs as a continuous predictor variable. This analysis revealed no significant main effect for Time 1 personal fixed beliefs ($F(1,32) = .005, ns$), but the interaction effect was still significant ($F(1,32) = 4.16, p < .05$). This indicated that the group norm, in interaction with group identification, predicted significant reductions in Social Interaction Anxiety, over and above any possible effects of participants' personal beliefs about the malleability of their own shyness.

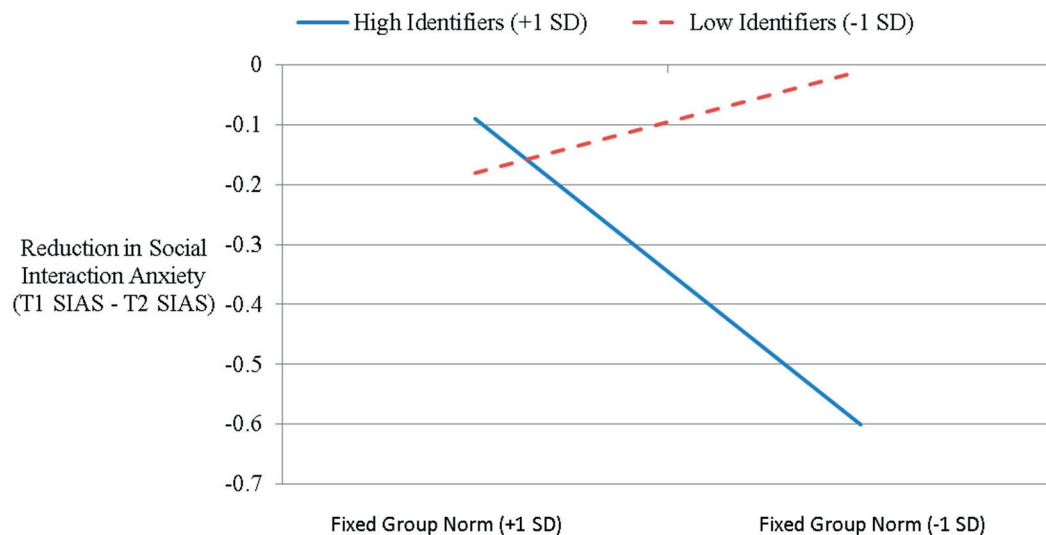


Figure 11: Shy identification and a fixed group norm (at Time 1), interact to predict reductions in Social Interaction Anxiety at Time 2.

A final analysis was run to assess whether Time 1 identification with ANU students predicted reductions in shyness—either as a main effect, or in interaction with beliefs about the group norm. This analysis was run on the basis that identification with the group of shy first-year ANU students was quite highly correlated with identification with the overarching category “ANU students”. This raised an alternative explanation of the results, namely that it

was not identification with the group of shy first-year ANU students *specifically* that was driving the effect, but merely identification with ANU students as a wider category. Also, as briefly noted earlier, Valentiner et al. (2011) found that “college belongingness” appeared to be associated with lower shyness levels over the course of the transition to university. To assess this possibility, a within-subjects ANOVA was again run with Time 1 and Time 2 SIAS as the within-subjects variable, and ANU identification, the group norm, and their interaction, as between subjects, continuous predictor variables. This analysis revealed no significant main ($F(1,33) = , ns$) or interaction effect ($F(1,33) = , ns$). This suggested that the pattern of findings could not be explained as an outcome of identification with the wider group of ANU students.

Study 4 discussion

The results supported the hypothesis that identification with a group of fellow, shy students predicted significant reductions in social interaction anxiety over the course of the transition to university, but only to the extent that participants perceived there was a group norm that shyness can change. This was clearly consistent with Empirical Proposition 4: that disorder identification, in interaction with a norm of malleability, will predict reductions in symptomatology over time. Interestingly, there was no evidence that personal malleability beliefs had any independent main effects in predicting reductions in social interaction anxiety in the transition to university. This contrasts with the findings of Valentiner et al. (2011) who found that those university students with high baseline levels of shyness, but who personally believed that their own shyness levels were relatively malleable, experienced lower subsequent levels of social performance anxiety over the course of the year.

This lack of an effect associated with Time 1 personal shyness malleability beliefs may in part be due to the disappointing response rate obtained in the current study and the consequently low statistical power. While the response rate of 24% obtained in the present

study was similar to the 22% obtained in the study by Valentiner et al. (2011), owing to the relatively small baseline sample in the current study ($N = 156$) the final sample was small and so the findings here obviously need to be interpreted cautiously.

Nevertheless, the findings are theoretically entirely consistent with the central theoretical proposition that psychologically identifying with disorder categories need not be a barrier to change. Importantly, the present findings begin to clarify how these group processes can have substantive and enduring effects on disorder symptoms. By contrast, much research examining the role of group identities on behaviour, particularly in the stereotype threat and cognitive priming literature, are highly situational and transitory phenomena (Bargh & Chartrand, 1999; Steele & Aronson, 1995).

It was also notable that merely identifying with the superordinate category of ANU students, in interaction with a perceived malleability norm of change among the group of “shy first-year ANU students”, did not predict significant symptom changes. This finding was consistent with Empirical Proposition 3, tested explicitly in Experiment 3, that there should be some congruence between the defining features of the self-category and the normative beliefs and shared theoretical understandings of that category.

The current findings also point to some interesting directions for future research. A further area of investigation would be to examine in more detail how beliefs about shyness and malleability, at the personal and group level, dynamically evolve over time. Again, this was well beyond the scope of the current analysis, but the findings point to the interesting possibilities of examining these beliefs in the context of transitions such as starting university. The ideal study design would involve a longitudinal, multi-level model assessing how group level variation in the malleability norm, in interaction with identification, predict changes in symptomatology. Such a design would address an important limitation of the current study in which individual perceptions of the group norm were used as a proxy for the

actual shared beliefs of the group. A multi-level longitudinal model, by contrast, would make it possible to directly measure group level beliefs about the malleability of shyness and therefore provide more compelling evidence that it was indeed shared, group level beliefs, rather than individual representations of the group, that were causally relevant.

Such multi-level data would also provide the most rigorous evidence that collective mobility can reverse the negative spiral that often characterises anxiety disorders. Storms and McCaul (1976) coined the term “self-exacerbating syndrome” to describe certain conditions, such as insomnia, procrastination, and various anxiety conditions, because the experience of symptoms in moderation further exacerbates those symptoms (cf., Ascher & Schotte, 1999; Wegner, 1997; Wegner, Broome, & Blumberg, 1997). This self-exacerbation process is also implicit in certain HPC models of mental disorder described in Chapter 1 of this thesis (for example, the process by which low mood leads to rumination, sleep disturbance, and further low mood; Wichers, 2014).

But the corollary of this self-exacerbation process is that incremental progress should dampen, and perhaps reverse, this negative spiral. The collective mobility thesis implies that, as fellow group members experience incremental, positive changes, this provides evidence and inspiration for others who, at an individual level, may be experiencing setbacks. But crucially, this buffering effect of group progress on individually experienced setbacks would be predicted to diminish if some group members change too fast (and thereby diminish their comparative fit), or if subsets of group members experience repeated setbacks, leading them to question either the *validity* of the normative theory, or the *applicability* of the theory to themselves.

This highlights the more prosaic point that although Study 4 demonstrated that shy group identification, in interaction with the norm of malleability, predicted reductions in shyness over a seven month period, the study was not designed to test the proposed

mechanism of change. Specifically, it is important to demonstrate that the malleability norm by identification interaction predicts more malleable personal beliefs and that this, in turn, predicts symptom reductions (i.e. Empirical Proposition 5). This proposition, in the context of a slightly different analogue disorder, was tested in Study 5.

Chapter 8: Study 5: The Interaction Between Disorder Identity and the Group Malleability

Norm Predicts Personal Malleability Beliefs and Lower Anxiety Symptoms.

In Study 5, I aimed to synthesise the findings of Studies 1-4, by empirically testing the basic theoretical model implied by the collective mobility thesis. I sought not only to replicate the finding that the interaction between identification and group malleability norms predicts individual malleability beliefs (Corollary 2, tested in Study 1) and reductions in symptomatology (Study 4), but examine how these two sets of findings are related. In doing so, I assess Empirical Proposition 4: the interaction between a malleable group norm and disorder identification will predict more malleable personal beliefs about the disorder which will, in turn, positively mediate the relationship between the malleable group norm and less symptomatic behaviour (i.e. conditional mediation, Hayes, 2011).

In addition to showing the mediational process underpinning the collective mobility thesis, it was also important, if possible, to test whether these proposed processes lead to observable behavioural changes. Although the self-report measures used in the previous studies provide rich insights into the proposed cognitive and emotional processes, there is an obvious concern that symptom self-reports (measured in Study 4) may be somehow confounded with measures of malleability beliefs. More generally, there is the concern that more positive symptom self-reports may not translate into objective behavioural changes.

However, it is also difficult methodologically to operationalise observable behavioural change in the context of the disorder analogue of shyness, used in the earlier studies. Although phenomenologically the anxiety associated with shyness can be intense, the observable symptoms of shyness can be quite diffuse and difficult to observe. Indeed Henderson and Zimbardo (1998, Zimbardo, 1977) describe the phenomenon of “shy extroverts”: people who behave in outwardly sociable and extroverted manner but are inwardly anxious and fearful of other people’s social judgement. This raised significant

concerns about the viability of shyness as an appropriate analogue disorder for testing behaviourally observable outcomes.

Given these concerns, a potentially more suitable alternative analogue disorder, which is distinct from (but closely related to) shyness (Blöte, Kint, Miers, & Westenberg, 2009), is speech anxiety. The phenomenon of speech anxiety, also known as public speaking anxiety, is very commonly experienced in both healthy and clinical populations, with public speaking being among the most feared situations that people experience (Pull, 2012). Previous research has also demonstrated that malleability beliefs have significant impacts on emotion regulation while delivering a short speech (Kneeland, Nolen-Hoeksema, Dovidio, & Gruber, 2016b). Although Kneeland et al. did not find direct experimental evidence that their intervention decreased state anxiety, other researchers have found that the increased use of emotion regulation strategies decreases both self-reported avoidance behaviours (Kappes & Schikowski, 2013) and objective measures of anxiety while delivering a speech (Jamieson, Nock, & Mendes, 2012). On balance, then, speech anxiety appeared to be well-suited as an analogue disorder for deriving a specific, behavioural outcome measure that can be used to assess the collective mobility thesis in the context of anxiety disorders.

The research by Kneeland et al. (2016b) in the context of speech anxiety provides important insights into a rapidly growing literature investigating the cognitive mechanisms linking malleability beliefs to clinical outcomes in general (see Kneeland et al., 2016a for a review) and anxiety disorders in particular (Schroder et al., 2016). However, as noted in Chapter 5, this research provides a very limited analysis of the social origins and group dynamics of malleability beliefs. One of the very few studies to examine how social identity processes, and normative beliefs about change, can influence clinical outcomes, is provided by a study of recovering substance addicts (Buckingham, Frings, & Albery, 2013). Although in a different domain, the researchers found that identity differentiation (which the authors

defined as the difference between their level of identification as a “recovering addict” and an “addict”) was correlated with perceived self-efficacy and reductions in self-reported symptom behaviour. In a second study, these same authors found similar results with respect to recovering smokers. By measuring the perceived difference between the groups “recovering addicts” and “addicts” the authors indirectly captured the content of a shared group membership that endorsed the notion of recovery but *not* addiction. The authors note that this is consistent with the central notion within Self-Categorization Theory: that the normative content of a group is partly defined by social comparison processes (an understanding of who we are in contrast to who we are not; Oakes, 1987). Accordingly, the authors argue that their findings support the idea that individuals who displayed such evaluative differentiation would be inclined to conform to the norms associated with recovery. And, indeed, these findings are quite consistent with the collective mobility thesis.

However, while the findings by Buckingham et al. (2013) provide an important insight into the application of these ideas in the context of symptom recovery, the use of a difference score to operationalise the group norm arguably obscures the normative influence processes that underpin the observed behavioural changes. It is unclear, for example, which self-categorization the recovery norm relates *to*, since participants who rate their addict and recovering addict identities positively and neutrally, respectively, are treated equivalently to those who rate these identities as neutral and negative. A clearer demonstration of these ideas would, therefore, ideally include a more direct measure of the group norm.

Moreover, Buckingham et al.’s (2013) analysis can be extended, in line with the collective mobility thesis, by arguing that the strength of identification with a salient identity and the recovery norm should *interactively* predict an increase in the level of normative behaviour. According to the Social Identity Approach, only to the extent that a particular group membership is subjectively meaningful to the self-concept, and psychologically salient

in a particular context, are the associated group norms likely to shape behaviour (Haslam, Turner, Oakes, Reynolds, & Doosje, 2002).

The present experiment therefore sought to build on these initial insights by Buckingham et al. (2013) and Kneeland et al. (2016b) in the context of the sub-clinical condition of speech anxiety. Importantly, the study extends the research by Buckingham et al. by manipulating the representation of the group norm of malleability, and thereby provides stronger evidence that norms are causally related to personal malleability beliefs and behavioural outcomes. Specifically, it is hypothesised that the manipulation of a group malleability norm should influence individual malleability beliefs, but only to the extent that individuals identify with the group (consistent with Corollary 2). In turn, it is predicted that people's beliefs that speech anxiety is malleable should be negatively related to observable anxiety behaviours when they deliver a short speech (consistent with Empirical Proposition 4).

Study 5 method

Participants and design

One-hundred-and-two third-year undergraduate psychology students from an Australian National University course (median age = 20 years; 80 females) completed the study. Six laboratory sessions were run over the course of one week, varying in size from 16–18 students. Participants in each session were presented with one of two experimental conditions for the main manipulated variable (group norm: fixed, malleable) and answered questions assessing the main, measured predictor variable of social identification with the group of people who experience speech anxiety. The two experimental conditions involved participants interpreting half-page “word clouds” that were very similar to those used in Studies 1–3. The laboratory setup required participants to sit next to each other. Because participants might have noticed differences in the “word cloud” (which were quite large) the

two conditions were randomised across the six sessions, rather than across individuals, thus eliminating the risk that participants in one condition may be inadvertently exposed to the other experimental condition.

Materials and procedure

Phase 1: introduction to the study and baseline measures. After entering the laboratory, the experimenter explained to participants that they would first answer some questions about speech anxiety and then deliver a short speech before answering some final questions.

Participants each received a questionnaire that they completed individually. Six items were drawn from the public speaking scale of the Personal Report of Communication Apprehension (PRCA-24) (McCroskey, 1978), using a 1 (“Strongly Disagree”) to 7 (“Strongly Agree”) Likert scale to measure participants’ overall levels of public speaking anxiety. Participants then rated their individual implicit beliefs about the relative malleability of their speech anxiety using five items adapted from Beer’s (2002) study of shyness (e.g., “I have a certain amount of speech anxiety and it’s something I can’t do much about”).

Phase 2: group norm manipulation and social identification measure. After completing the above 12 items, participants were presented with a “word cloud” containing 18 words of various sizes either relating to the relatively fixed nature of anxiety (in the “fixed norm” condition; e.g. “can’t change”, “stuck”, “biological”), or the relatively malleable nature of anxiety (in the “malleable norm” condition; e.g. “can change”, “transition”, “can overcome”, “learn”). These word clouds were identical to those used in the earlier studies except that the words “speech anxious”, rather than “shy”, appeared in the middle of the cloud. Participants were told that the “word cloud” represented the words speech-anxious people used to describe their condition. It was explained that more frequently used words were in relatively larger text compared to less frequently occurring words. As a manipulation

check, participants then answered five items measuring the perceived group norm of malleability, again using modified items from Beer (2002; e.g. “Speech anxious people have a certain level of speech anxiety, and it is something that they can’t do much about”). Participants then rated the extent to which they agreed with the statement, “I identify with the group of speech-anxious people who contributed to this word-cloud”—based on a validated, single item measure of social identification (Postmes, Haslam, & Jans, 2012). This shorter measure was used, rather than the multi-item measure used in the earlier studies in part because of concerns that an overly long time delay between the main manipulation and the measurement of the dependent variable may attenuate the impact of the norm manipulation.

Phase 3: forming pairs and delivering a speech. Participants were then asked to pair with another participant with whom they would deliver a three minute speech explaining the prisoner’s dilemma. The topic of the speech was chosen because it was relevant to participants’ coursework, and was relatively complex and so might be expected to elicit some anxiety.

Participants were randomly assigned to either be the first speaker or second speaker. First speakers were instructed to deliver the speech to their partner while standing up, and informed that their partner would be rating them on their observed anxiety levels. Immediately after delivering the speech, the first speaker again rated his or her own individual implicit beliefs scale assessed in phase 1. Participants also rated their state anxiety while delivering a speech using five items assessing state anxiety (Marteau & Beyyter, 1992), by indicating their level of agreement on a 7-point Likert scale using a 1 (“Strongly Disagree”) to 7 (“Strongly Agree”) scale to the statements: “During the speech I felt calm (reverse scored)/ tense/ relaxed (reverse scored)/ content (reverse scored)/ worried”.

While the first speaker was answering these items, the second speaker (the observer) rated in private on a scale of 0 (totally relaxed) to 100 (severely anxious) his or her partner’s

“highest level of anxiety that he or she experienced during the speech”. After completing these ratings, partners then swapped roles and followed the same procedure described above. Participants then received a full debriefing.

Study 5 results

Measures and manipulation

Public Speaking Apprehension Scale. The six items assessing public speaking apprehension (i.e., anxiety) formed a reliable scale (Cronbach's $\alpha = .90$). The mean anxiety score was significantly higher than the mid-point of the scale ($M = 4.65$, $SD = 1.35$; $t(101) = 4.88$, $p < .001$) indicating that, overall, participants experienced at least some speech anxiety symptomatology. Given that speech anxiety is not a clinical condition, and there are no validated guidelines for defining speech anxiety categorically, all participants were included in the subsequent analysis.

Individual Fixed Beliefs. A series of principal components analyses with varimax rotation were conducted to assess the unidimensionality of the multi-item measures of implicit beliefs about speech anxiety. The six items assessing participants' Time 1 (baseline) individual implicit beliefs loaded on two components with an eigenvalue over 1; the first explained 45.80% of the variance, and the second explained 25.80% of the variance. The same Time 2 individual implicit beliefs measure also loaded on two components, explaining 59.80% and 17.91 % of variance. The two components corresponded with items assessing fixed and malleable beliefs respectively. Visual inspection of the scree plot for the Time 1 and Time 2 measures suggested that the Time 1 measure had two main components, whereas the Time 2 measure displayed only a single component. Unlike the earlier studies, this suggested that this measure, in the context of speech anxiety, was not clearly unidimensional. In these circumstances, other researchers have simply used the two items from the “fixed belief” subscale because this can be more reliable (Hong et al., 1999). An individual “fixed

belief” measure was, therefore, derived by computing the mean of the two items: “I have a certain amount of anxiety and it’s something I can’t do much about” and “My speech anxiety levels are something about me that I can’t change very much”.

This measure displayed adequate reliability at Time 1 (Cronbach’s $\alpha = .86$) and Time 2 (Cronbach’s $\alpha = .86$) and was used in the subsequent analyses. The two similarly worded items within the implicit group norm scale (e.g. “Speech-anxious people have a certain amount of speech anxiety and it’s something they can’t do much about”) also displayed adequate reliability (Cronbach’s $\alpha = .94$), and so the mean of these two items was used to measure the fixed group norm.

Anxiety measures: The five items assessing state anxiety formed a reliable scale (Cronbach’s $\alpha = .92$). Principal components analysis was conducted to assess the unidimensionality of this scale and this revealed a single component with an eigenvalue greater than 1 and which explained 76.04% of the variance in these items. Each individual’s state anxiety score was calculated by taking the mean of the five items (reverse scoring the 3 positively worded items). The raw score on the single-item anxiety rating by observers was used as the measure of observer rated anxiety.

Group Norm Manipulation: The fixed group norm measure was significantly higher in the fixed norm condition ($M = 5.29$, $SD = 1.69$) than the malleable norm condition ($M = 2.75$, $SD = 1.23$, $t(100) = 8.72$, $p < .001$), indicating that the manipulation was successful in changing the perceived norm regarding the malleability of speech anxiety.

Social Identification. Identification with the group of speech anxious people who contributed to the word cloud was hypothesised to moderate the influence of the group norm on participants’ individual malleability beliefs. However, as identification was measured after the manipulation, it was important to ensure the manipulation did not, in itself, influence participants’ identification levels. A between subjects t -test showed no significant

differences in identification across the two manipulated conditions ($p > .05$), and so measured identification was included as a moderator in the main analysis.

Main analysis

The means, standard deviations and bivariate correlations between the main study variables are displayed in Table 12. To test the hypothesis that the interaction between the group norm and identification predicted individual fixed beliefs, Time 2 individual fixed beliefs were regressed on the manipulated group norm, (mean-centred) identification, and the norm by identification interaction effect. This revealed significant effects for the norm ($t(100) = 2.23, p = .03$), identification ($t(100) = 3.06, p < .01$), and their interaction ($t(100) = 2.47, p = .01$). Higher identifiers (1 SD above the mean) displayed significantly higher fixed beliefs in the fixed norm condition ($M = 4.65$) than in the malleable norm condition ($M = 3.23, t(100) = 3.32, p < .01$), whereas for lower identifiers there were no differences between the fixed norm ($M = 2.97$) and malleable norm conditions ($M = 3.05, t(100) = -.17, ns$).

Table 12: *Study 5 means, standard deviations and bivariate correlations between main study variables*

Measure	M	SD	1.	2.	3.	4.	5.	6.
1. Public Speaking Apprehension	4.65	1.34	--	.58**	.48**	.52**	.36**	0.58**
2. Individual Fixed Belief (T1)	3.34	1.52		--	.42**	.70**	.27**	0.32**
3. Identification	3.45	1.66			--	.27*	.02	0.12
4. Individual Fixed Belief (T2)	3.40	1.57				--	.25**	0.46**
5. Observer Rated Anxiety	37.64	24.23					--	0.47**
6. Self-rated state anxiety	4.15	1.53						--

* Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

The foregoing analysis is consistent with Study 1 (Corollary 2) in showing that the malleability norm by identification interaction predicts individual malleability beliefs. However, it was noted that, unlike in the earlier studies, identification was measured slightly differently in the sense that participants were asked to rate their agreement with the statement “I identify with the group of speech-anxious people who contributed to this word cloud” rather than just the statement “I identify with the group of speech-anxious people”. This raises the concern that prior individual fixed beliefs may confound the interpretation. It may be, for example, that individuals identified with the group because it confirmed their prior beliefs about the malleability of their anxiety, rather than because group identification *per se* moderated the effect of the group norm manipulation. To test this alternative explanation for

the results, the analysis was repeated, controlling for Time 1 individual fixed beliefs (and its interaction with the group norm, as recommended by Yzerbyt, Muller and Judd (2004)), to assess the independent effect of the norm by identification interaction effect on participants' Time 2 individual fixed beliefs. This revealed a significant main effect for prior individual fixed beliefs ($t(100) = 7.96, p < .001$) and the predicted significant interaction between the group norm and identification ($t(100) = 2.05, p < .05$). No other main effects or interactions were significant. This analysis confirmed that participants' prior individual fixed beliefs could *not* explain the significant norm by identification interaction effect on participants' Time 2 individual fixed beliefs.

To probe the interaction, simple slopes analysis was conducted, controlling for Time 1 individual fixed beliefs (and its interaction with identification), by evaluating the effect of the norm manipulation on Time 2 individual fixed beliefs at 1 SD above and below the mean for identification. There was a significant slope for high identifiers ($t(100) = 2.15, p = .03$), but no significant slope for low identifiers ($t(100) = -.87, p > .1$). The pattern of slopes was consistent with the hypothesis that, relative to low identifiers, high identifiers displayed more malleable beliefs when presented with a malleable norm and more fixed beliefs when presented with a fixed norm.

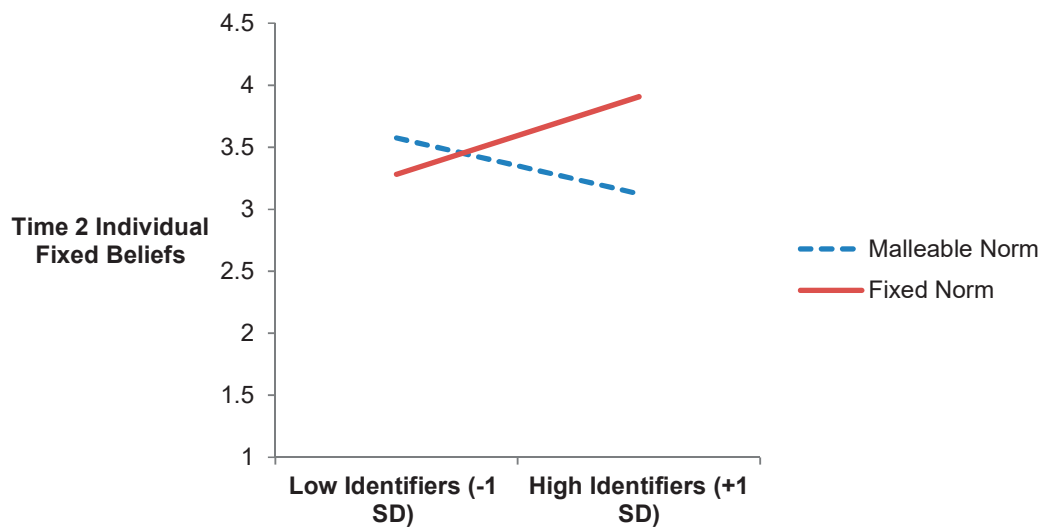


Figure 11: Significant interaction between the malleability norm and group identification on Time 2 individual fixed beliefs, controlling for prior (Time 1) individual fixed beliefs and the interaction between the norm and prior individual malleability beliefs.

Moderated mediation analysis

Table 14 shows that Time 2 individual fixed beliefs were significantly related to self-reported state anxiety and observer ratings of anxiety. But to more rigorously test for moderated mediation, it was necessary to assess more formally the evidence for a conditional indirect effect in which the norm by identification interaction effect influenced Time 2 individual fixed beliefs and then, in turn, influenced participants' self-reported and observed anxiety levels. Again, it was important to control for both Time 1 individual fixed beliefs, and its interaction with measured identification.

The proposed moderated mediation model was tested using Hayes's (2012) Process Package (Model 9). The model partly replicated the second analysis described above, but treated Time 2 individual fixed beliefs as a mediator, predicting, firstly, self-reported anxiety levels, as shown in Figure 12. Following the recommendations set out by Muller, Judd, and Yzerbyt (2005), moderated mediation was assessed by testing whether the moderated effect

between the norm manipulation and individual Time 2 fixed beliefs was significant, and whether the direct effect of individual Time 2 fixed beliefs predicted self-reported anxiety. There is evidence for moderated mediation if there is a significant interaction effect predicting the mediator and a significant effect of the mediator on the outcome variable. As shown in Figure 12, the norm by identification interaction effect significantly predicted more fixed beliefs, which in turn predicted self-reported state anxiety. This confirmed that individual fixed beliefs mediated the relationship between the group norm and self-reported state anxiety, conditional on participants' identification.

To assess moderation of the indirect path, bias corrected bootstrapping was used to generate 95% confidence intervals for the direct and indirect effects. If confidence intervals do not include zero, then this indicates a statistically significant effect. Results were based on 10 000 bootstrap samples. This analysis revealed that for high identifiers (1 SD above the mean) the confidence interval for the indirect effect did include zero ($IE = .34$, lower 95% $CI = -.06$, upper 95% $CI = .90$), and for low identifiers the indirect effect did include zero ($IE = -.26$, lower 95% $CI = -4.86$, upper 95% $CI = 2.73$). Thus, although both effects tended in the opposite directions at high and low identification, neither of the indirect effects was statistically significant. There was also no significant direct effect between the norm manipulation and anxiety levels ($DE = 6.83$, lower 95% $CI = -2.53$, upper 95% $CI = 16.20$). Although the indirect effect did not reach significance for high identifiers, as noted by Muller et al. (2005, see also, Hayes, 2015), this is not a precondition for demonstrating moderated mediation.

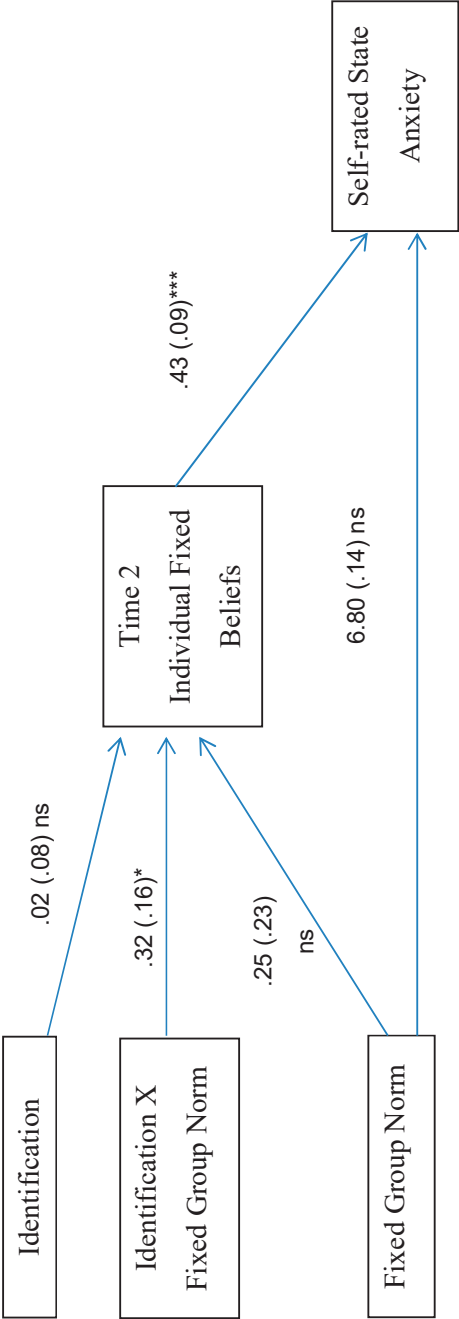
Having confirmed the proposed moderated mediation model for self-reported state anxiety, the same analysis was performed for observer rated anxiety. Before doing this, however, it was important to check whether the fact that participants were paired up when delivering and rating each other's speech may have influenced their anxiety ratings which

would then require the use of multilevel modelling techniques to control for interdependence between raters. This risk should only be relevant to the partner anxiety ratings, since there is no reason to suspect that partner effects might emerge prior to participants actually interacting and, overall, the risk of ratings influencing each other should be low since participants completed all ratings privately. However, to test and adjust for possible interaction effects creating non-independence between partner anxiety rating scores, the recommendations outlined by Kenny, Kashy and Bolger (1998) for the analysis of dyadic data were followed. This involved assessing the partial Pearson product moment correlation coefficient between the first and second speaker ratings of partner anxiety levels. A non-significant correlation indicates that the ratings can be treated as independent observations. This analysis revealed no statistically significant relationship between partner anxiety ratings ($r = .20, p = .16$), and so, as recommended by Kenny et al. (1998) partner anxiety ratings were treated as independent observations.

The same moderated mediation model, using observer rated anxiety as the outcome variable, was then tested and substantively similar results were obtained, as shown in Figure 13. The significant interaction effect between identification and group norm on Time 2 fixed beliefs, and significant effect of Time 2 fixed beliefs on observer ratings of anxiety provided evidence for moderated mediation. Again, however, although the coefficients of the indirect paths tended in opposite directions at one standard deviation above ($IE = 2.73$, lower 95% $CI = -.42$, upper 95% $CI = 8.00$), and below ($IE = -1.02$, lower 95% $CI = -4.73$, upper 95% $CI = 1.23$) the mean of social identification, neither of these indirect paths reached statistical significance.

In sum, despite not finding any significant indirect paths (when evaluated at 1 SD above and below the mean of identification), the two conditional mediation analyses confirmed the hypothesis that identification moderated the relationship between the fixed

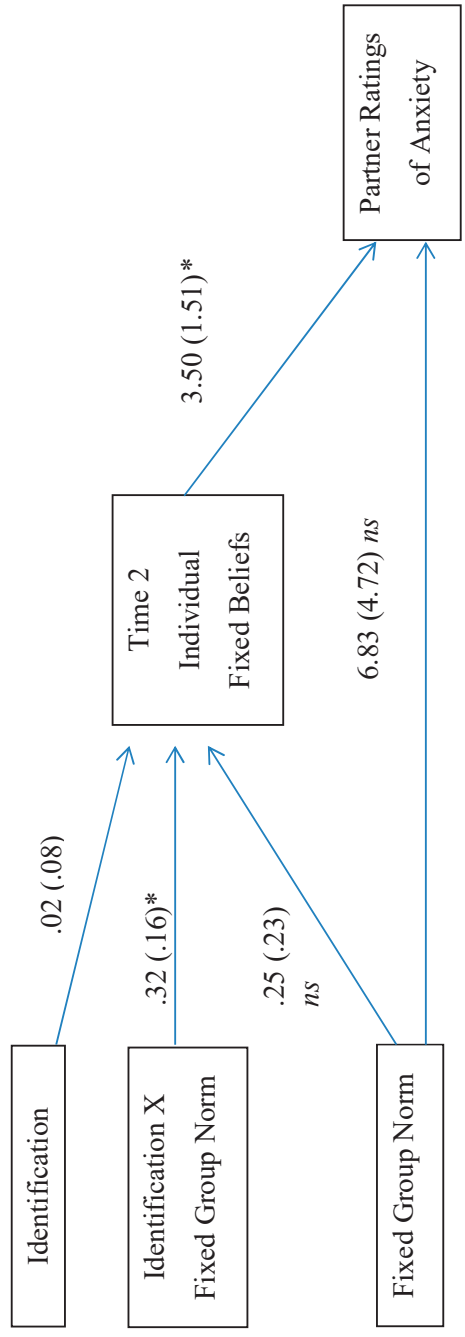
group norm and individual fixed beliefs at Time 2, and these individual fixed beliefs further predicted participants' self-reported state anxiety and observed anxiety while delivering a speech.



Key: Unstandardised regression coefficients are displayed with standard errors in brackets. For visual clarity, the path from Time 1 Fixed beliefs to Time 2 Fixed beliefs ($b = .68$, $SE = .09$, $p < .001$) and the path from interaction between Identification and Time 1 Fixed beliefs to Time 2 Fixed beliefs ($b = -.20$, $SE = .17$, ns) are not displayed.

* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 12: Conditional mediation model assessing moderation of the fixed group norm by group identification on Time 2 Individual Fixed Beliefs and subsequent relationship to Self-rated State Anxiety (controlling for the effect of Time 1 Fixed Beliefs and the interaction between Time 1 Fixed Beliefs and Identification).



Key:

Unstandardised regression coefficients are displayed with standard errors in brackets. For visual clarity, the path from Time 1 Fixed beliefs to Time 2 Fixed beliefs ($b = .68$, $SE = .09$, $p < .001$) and the path from interaction between Identification and Time 1 Fixed beliefs to Time 2 Fixed beliefs ($b = -.20$, $SE = .17$, ns) are not displayed.

* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 13: Conditional mediation model assessing moderation of the fixed group norm by group identification on Time 2 Individual Fixed Beliefs and subsequent relationship to Partner Ratings of Anxiety (controlling for the effect of Time 1 Fixed Beliefs and the interaction between Time 1 Fixed Beliefs and Identification).

Study 5 discussion

This study tested whether identifying with a group of fellow sufferers of speech anxiety may help group members to overcome their symptoms, depending on the group's normative understanding about the malleability of the condition. The results showed a pattern of moderation that was consistent with Corollary 2, and thereby replicated the findings of Study 2. Specifically, individuals who identified with speech anxiety sufferers, and were informed that fellow speech anxiety sufferers held a group norm that speech anxiety is relatively malleable, reported that their personal speech anxiety was more malleable. By contrast, low identifiers were not significantly influenced by the norm. Further, and consistent with Empirical Proposition 5, moderated mediation analysis showed that the relationship between a fixed group norm and self-reported and observer-rated anxiety, via personal fixed beliefs, was more positive for high identifiers compared to low identifiers. These findings provide support for the collective mobility thesis by showing that individual malleability beliefs and subsequent symptomatic behaviours can be an outcome of the interaction between disorder identification and the group malleability norm.

The current study therefore provides important evidence that links together the key findings from Studies 1-3, showing the impacts of malleability norms and disorder identification on individual beliefs, and the findings from Study 4, showing how malleability norms influence symptom outcomes. The findings of the current study provide a conceptual replication and extension of these previous results. In particular the current study addresses a key limitation of Study 4 by showing that the proposed collective mobility thesis provides a basis for predicting not just self-reported, but observer-rated, behavioural outcomes, while demonstrating how individual malleability beliefs mediate the relationship between group processes and symptoms.

There are also limitations to this study. Firstly, it would have been preferable to be able to take video recordings of participants while they delivered the speech, and to have these rated by observers completely independent of the study, in order to eliminate possible bias in the partner anxiety ratings. Although there was no evidence of such bias statistically, it would have been better to address this issue through experimental design. Secondly, identification was a measured variable and so it is not possible to rule out alternative, non-causal, hypotheses for explaining the observed pattern of results. However, it should be noted that the fact that the group norm was manipulated, rather than measured, provides greater confidence that the association between group norm and individual malleability beliefs do not merely reflect common method variance, or demand characteristics. Thirdly, as with Studies 1-3, the advantages of experimentally manipulating the norm in a controlled laboratory environment, comes with the obvious trade-off of more limited ecological validity.

Despite these limitations, the findings are a small step towards demonstrating the empirical viability of the overarching Disorder Identity analysis, developed in the theoretical chapters of this thesis. To reiterate, the theoretical argument was that Disorder Identities provide the mechanism through which shared ideas about the nature of a disorder get in under the skin and change how individuals understand themselves and that this in turn changes how people with disorder behave and experience symptoms. The challenge was to empirically operationalise this idea, using the theoretical tools developed within Social Identity Theory and Self Categorization Theory. The present findings suggest that these tools can indeed provide a basis for scrutinising the Disorder Identity analysis and, by extension, the claims implied by the “looping kinds” argument (Hacking, 1999). But this enthusiasm needs to be tempered with the acknowledgement that the research agenda implied by the Disorder

Identity analysis would require an array of quantitative and qualitative methodologies at varying levels of analysis.

It is worth noting the implications of this point in the context of the evidence for the collective mobility thesis provided by Study 4 and Study 5. In Chapter 5 I argued that the collective mobility strategy is likely to have both positively heterodynamic consequences (decreasing the numbers of those defined, categorically, as “disordered”) and negatively homeodynamic consequences (decreasing the severity of symptoms experienced by sufferers). The present study, of course, was not designed to be able to distinguish between these two outcomes. For a start, speech anxiety (as with the study of shyness) is merely an analogue condition that is very widely experienced. Consequently, there are no institutionally defined “cut points” for deciding when speech anxiety is problematic (in a way that such institutional decisions must be made for clinical conditions). The point of the present experiment was merely to show, in a theoretically coherent manner, how the proposed collective mobility process could reduce symptoms. To the extent speech anxiety is a valid analogue for a diagnosable condition, such as social phobia, then it would be expected that these sort of reductions in symptomatology, if sustained and on a large enough scale, would decrease the numbers of people who meet criteria for a disorder as well as the severity of symptoms experienced by those with disorder. Empirically testing this sort of macro-level change would obviously require large, population level longitudinal surveys assessing, among other things, disorder identification(s), normative understandings of disorders, and including diagnostically validated measures of symptomatology.

Chapter 9: Conclusion

The central span on the bridge from physics to society is collective intentionality, and the decisive movement on that bridge in the creation of social reality is the collective intentional imposition of function on entities that cannot perform those functions without that imposition. (Searle, 1999)

This thesis can be understood as an extended elaboration, defence and application of this astonishing claim that collective intentionality (shared beliefs) can systematically change the way that humans think, feel and behave, and in a way that is not reducible to the intrinsic properties of them in isolation. Specifically, I have argued that what Searle calls the “collective intentional imposition of function on entities” is a particular kind of categorization. In the case of mental disorder, it is an act of social categorization that explains how disorder can simultaneously be a part of both physical and social reality. These mental disorder social categories take the form of HPCs, which are operationalised by experts, and as *self*-categories, defined from the perspective of those who perceive themselves to be a part of a disorder-related group. I have further argued that analysing mental disorder in this way offers novel and theoretically generative insights into the nature of disorder and provides a framework that reveals the extraordinarily diverse ways that people with disorder can “recover”. Perhaps most importantly from a clinical perspective, I have derived, and found empirical support for, a central prediction from this framework by showing that Disorder Identities need not necessarily be a barrier to change, but could, in some circumstances, help individuals to overcome their symptoms, depending on the group’s normative beliefs about the malleability of the condition.

In this final chapter, I start by reviewing the pivotal twists and turns that have led from the crisis of categorization within psychiatry discussed in Chapter 1, to the Disorder Identity Analysis of recovery in Chapter 2, the social-psychological processes underlying this analysis of recovery as discussed in Chapters 3 and 4, and 5 and finally the empirical analyses assessing the collective mobility thesis in Chapters 6 to 8. I then summarise some of the limitations, implications and opportunities of the overarching analysis as they relate to clinical practise, Social Identity theorising, and theoretical understandings of disorder more generally.

Framing the issue: What is mental disorder and where is it?

This thesis began with a conceptual problem: what explains coherence, stability and change observed in mental disorder across time and space? One popular answer to this question is that mental disorders are diseases like diabetes or HIV, and so conceptual variation in mental illness is no more mysterious than the fact that these biomedical conditions vary across social and environmental contexts. The problem with this answer is that, despite extraordinary medical advances, and unlike any other clinical specialty, it would seem that there are no specific underlying entities, or patho-aetiological processes that can explain any mental disorder. This has led to a crisis within psychiatry. Importantly, it is not just the usual suspects within critical theory (Foucault, 1964) and anti-psychiatry (Laing et al., 1970; Scheff, 1970; Szasz, 1961) saying this. As Zachar and Kendler (2007) note, there are “no more spirochetes to be found” (p. 563). Allan Frances, the former chair of the DSM-IV taskforce, is more blunt: “mental disorders don’t really live ‘out there’ waiting to be explained. They are constructs we have made up—and often not very compelling ones at that.” Some hard biomedical reductionists propose redoubling efforts to find the neurobiological underpinnings of mental disorder (Insel et al., 2010), but many fear there is

something fundamentally conceptually wrong with the disease model (Kingdon & Young, 2007). But if the disease model is to be overturned, then what should replace it?

I have argued that Boyd's (1991, 1999) HPC model provides a compelling alternative to the disease model because it articulates a non-essentialist account of scientific categories that can be applied to mental disorder. The HPC account is not anti-realist and does not permit utterly unrelated things to be lumped together on the basis of superficial resemblances. It hangs on to objective things in the world (mechanisms) that explain how certain surface level properties tend to co-occur. But crucially, and as with constraint relations models of categorization more generally, the account also acknowledges that categories also depend, in part, on their usefulness in predicting and explaining things of interest to a discipline. What are these "things of interest" to the disciplines of psychiatry and clinical psychology? Abnormality? Although true, this merely pushes the question back one step. Disordered behaviours and brain circuits do not come pre-packaged with little stickers declaring their abnormality. No, we *write* those stickers. But who are "we", where do "we" come from and who do "we" represent? Mostly, psychiatrists and clinical psychologists do not believe they come fully-formed from a plane of eternal truth, bearing unquestionable insight and authority. Rather, they recognise these are special groups of people who society has delegated the fraught and critically important job of using science to operationalise wider social judgements about who is and who is not normal. The existence of these disciplines is therefore deeply dependent on society. Firstly, society defines the very object of study, in part because, without knowing what normal looks like, the word abnormal makes no sense. Secondly, it is society that underwrites the institutional power of these disciplines. This leadership role can be understood as a "special gift" bestowed upon experts for being representative of "us", as a wider community (Platow et al., 2006).

These claims seem radical. Actually, the basic point is a mundane consequence of a centuries old insight: one cannot derive claims about how people *should* be from descriptions of how they *are* (Hume, 2012 / 1739). The DSM, since version III at least, is utterly explicit that definitions of disorder require clinicians to make normative judgements (Cooper, 2015). Moreover, when pushed, even most biological psychiatrists concede that this “normativity problem” is real and unavoidable (Brulde, 2007), however begrudgingly (Walter, 2013). It is perhaps surprising then, that the HPC model, which in many ways makes more transparent these uncomfortable truths, has nevertheless been enthusiastically received as a potential circuit breaker for the crisis within psychiatry (Kendler et al., 2011). In part, this is because the account’s emphasis on empirically discoverable mechanisms helps restore the metaphysical credibility of psychiatric disorder, and shows that, in fact, many of the problems with essentialism are problems in other natural sciences too.

But despite this important conceptual move away from essences and towards mechanisms, much of the conceptual baggage of the disease model has been retained. In particular, I have argued that the question of *where* is mental illness weighs heavily on HPC accounts of disorder—certainly much more so than many seem to have noticed. Of course it is still light work to refute naive externalist accounts of mental disorder. But it is much tougher to confidently dismiss the idea that people’s collective awareness of their own disorder will not create a feedback loop that causally influences the properties of a disorder category. The nub of Hacking’s (1999) “looping kind” argument is not just that: (1) boundaries of disorder change with changing social values; nor merely that, (2) people behave differently when they are aware of their disorder label (cf. Gergen, 1985). Rather, the argument is that claims (1) and (2) are connected. People’s *collective* self-awareness leads them to change themselves and the communities around them in ways that force experts to

peel off old disorder stickers, write new ones, and change how, when and to whom they are applied. This is *not* to claim that there are no brute, indifferent mechanisms that explain stability and change in mental disorder. Rather, if Hacking's thesis is right, then some of the mechanisms governing how symptoms co-occur and change must spill out beyond the boundaries of people's heads. Mental disorder, on this view, must be understood as a dynamic system. The implications of this claim are potentially profound. In part, this is because Hacking's claims suggest the "normativity problem" could actually help provide a resolution to the problem of explaining apparent stability and change in mental disorder.

But of course, Hacking's claims could be misguided—either because looping effects are empirically implausible or conceptually flawed—or because his claims amount to nothing more than a provocative, but ultimately unfalsifiable, philosophical conjecture. I argued in Chapter 1 that given what is known about symptom perception, placebo effects, and sociogenic illness, looping effects are certainly empirically plausible. Moreover, there is nothing conceptually problematic with looping effects as HPC mechanisms, so long as the arbitrary stipulation that mechanisms must be internal is lifted (cf. Kendler et al., 2011). The deeper problem is empirically operationalising social mechanisms.

Indeed, many suppose that letting go of internalism leaves the empirical study of mental disorder on a conceptual high wire. The skull, or perhaps the skin, seems to mark the very point where the natural sciences stop and the more fragile and superficial, social "sciences" begin (Wilson, 2004). Consequently, a common, usually unstated, assumption is that allowing the fundamental causal mechanisms that constitute disorder to traverse beyond the neocortex seems to be a move in precisely the wrong direction, making whatever confusion there is about mental disorder much, much worse.

But this implies a peculiar metaphysical discontinuity between the causal processes within heads and those that occur between and outside them. On the contrary, the world described by physics is a seamless expanse of microphysical objects, causally interacting over time and space. Far from being less “real”, it seems that social facts about everything from money to mental asylums have remarkable causal consequences on neural states but these are nevertheless irreducible to the intrinsic properties of metal discs, buildings or individual brains. How is it possible that such indubitably real effects emerge from such allegedly superficial externalist goings on? I have argued, with good company (Bruner, 1957; Durkheim, 1895; Searle, 1995; Turner et al., 1987), that there is something metaphysically fundamental about shared beliefs, “we-beliefs”, that underpins these extraordinary causal consequences.

But if these shared beliefs play such a consequential role, then this puts a tremendous burden on social psychologists in trying to explain the cognitive underpinnings of these beliefs. At first glance, it would seem that social psychology is simply not up to the job. Far from being the sorts of things that could create the staggering richness and complexity of human social arrangements, socially shared representations of groups of people, commonly known as stereotypes, have long been considered “to symbolise nearly all that is deficient in popular thinking” (Asch, 1952, p. 232). At best, stereotypes are crude and rigid simplifications. At worst they are dangerous distortions that are the root cause of stigma, prejudice, discrimination and disempowerment.

Despite this, an alternative view is that categorization is precisely what enables perception (Bruner, 1957) and that socially shared representations are fundamental to how people predict, explain and control themselves and others (McGarty et al., 2002). Informed by this alternative perspective, I have argued that SIT and SCT provide a coherent, and

empirically well-supported, account of the cognitive processes explaining how collective intentionality is possible.

Building the conceptual foundations of the disorder identity account

A central task of this thesis has therefore been to use the principles within SCT and SIT to develop the conceptual foundations needed to put externalist HPC disorder mechanisms on firmer ground as an object of empirical investigation. Even if one is unconvinced that externalist mechanisms have much to do with explaining stability and variation in disorder, at the very least by attempting to formulate the looping hypothesis empirically it might be possible to prove Hacking wrong. The pivotal move in laying these conceptual foundations has been to frame Hacking's "looping kind" argument as an intergroup process, underpinned by social categorization. The inspiration for this move has been to notice the parallel structure of HPCs (Boyd, 1991) and the structure of self-categories (McGarty, 1999): both are categorization models constrained by theory (mechanisms), perceived equivalence (similarity), and category use (disciplinary purpose). This makes it possible to refer to expert representations of disorder (from the third person perspective), and the *self*-representations of those with disorder, and to begin to explain how these representations might interact. This formulation enables a crucially important perspective switching. One can hold constant the definition of an HPC category and describe the consequences of identifying with disorder in different contexts. But one can also see how HPC categories, from the perspective of those with disorder, may be precisely what it is that group members seek to change, leading to conceptual "creep" in the scope of disorder categories (Haslam, 2016).

This leads to the second conceptual foundation, again described in Chapter 2. This relates to the question of what it is that groups of people who self-categorize as disordered are

using their disorder self-categories for. Whatever functions people who self-categorize as disordered are collectively assigning to disorder categories, it is often very different from the functions that groups of clinicians are assigning. The issue is further complicated by the fact that people can self-categorize as “disordered”, but not meet criteria for a disorder and conversely they can be collectively aware that others (usually experts or society more generally) perceive that they are disordered, but insist that their behaviour is actually functional or positively valenced.

To take account of these complexities, I defined the term “Disorder Identity” to refer to the set of people who hold the shared beliefs: that group members display(ed) a characteristic set of behaviours, that the set of behaviours is perceived as institutionally dysfunctional (even if group members disagree with this or disagree with the institutional claims about the causes of dysfunction), and that the group is subjectively meaningful. In turn, this definition provides the basis for what I called Disorder Identity Adaption Strategies that arise when an individual perceives that a Disorder Identity is contextually relevant to the self and the individual acts in a way that is informed by a shared understanding of what a disorder category means and the socio-structural constraints of the environment. Therefore, in much the same way that organisms carve out ecological niches, I argued that Disorder Identities are a resource that allows group members to modify themselves and their social environment.

The third conceptual foundation, described in Chapter 2, was to establish a more formal language and structure for describing the possible dynamic consequences of Disorder Identities on HPC disorder categories. Claiming that a phenomenon is dynamic is easy. Describing these dynamics precisely is much harder. While HPC accounts provide an important explanatory account of coherence and stability in disorder, this is arguably less

than half of what needs explaining. The looping kind argument implies that one of the reasons for this is that groups of people systematically change the properties of HPC disorders. But what kind of changes are these? To answer this question more precisely, I noted that HPC disorder categories seem not only to involve positively and negatively homeostatic mechanisms (that increase and decrease similarities between category members), but also homeodynamic mechanisms (that amplify or attenuate the number and severity of properties displayed by HPC category members), heterodynamic mechanisms (that facilitate the inflow and outflow of category members), and heterostatic mechanisms (that prevent inflow and outflow of category members). The advantage of elucidating these additional mechanisms is that it is not only possible to refer with more precision to HPC processes, but also allows researchers to concatenate these mechanisms to describe more complex sequential changes, and thereby provides the basic conceptual tools for describing the dynamics of HPC categories.

In sum, in Chapter 2, I established three important conceptual foundations of the Disorder Identity analysis by: (1) showing how mental disorder is an intergroup process, underpinned by social categorization; (2) defining Disorder Identity and Disorder Identity Adaption Strategies; and (3) systematically describing the set of mechanisms that underpin HPC dynamics. Along the way, I also showed how these three ideas pay immediate dividends by providing a clearer conceptual account of the notion of recovery-as-an-orientation. By contrast, extant theorising about recovery invariably becomes entangled by the normativity problem. The problems arise because researchers have typically attempted to provide an account of recovery that emphasises patient empowerment while at the same time insisting that certain kinds of collective behaviour on the part of consumer groups are somehow objectively disordered, irrational or dysfunctional. But this obscures the fact that,

whether or not these collective behaviours align with “our” norms, these actions are often highly subjectively meaningful from the perspective of those with disorder. An empirical account of these behaviours that neglects these collective meaning-making processes is therefore likely to be impoverished. By contrast, the Disorder Identity account I have developed aims to be more explicit about the normative judgements that underpin categorization processes and thereby reveals the very diverse ways that individuals can “recover”.

Social psychological mechanisms for explaining Disorder Identity dynamics

Although Chapters 1 and 2 develop important conceptual anchors, they provide few details about whether these postulated Disorder Identity dynamics are supported by, or even compatible with, extant social-psychological theory. In Chapter 3, I therefore reviewed various labelling theories and stereotype threat analyses that seem to pose a grave threat to the proposition that identifying with a disorder could, in any sense of the term, be considered adaptive. I argued that these counterpoints were far from decisive. Labelling Theory does not adequately account for the evidence that the stereotyping process is highly flexible, and context dependent; meanwhile modified labelling theory cannot adequately account for the so called “paradox of self-stigma”.

In Chapter 3 I also examined the literature showing the importance of lay understandings of whether putatively deviant behaviour is pathological, and if it is, then (in folk terms) whether it is in the brain, the mind, or the heart. I reviewed extensive evidence showing how lay theorising profoundly affects how people with mental disorder are treated by others, as well as people with mental disorder’s motivations to attempt to overcome their condition or to creatively adapt to it. Importantly, this literature is beginning to reveal the difficult trade-offs faced by people attempting to adapt to their disorder. In particular,

strategic essentialism seems to be effective in warding off some (but not all) kinds of stigmatising views, but only at the expense of prognostic pessimism (Kvaale et al., 2013). Moreover, these findings align with a growing literature showing that people's theories about whether a condition is relatively fixed or malleable determines their motivations to engage in therapy, to persist in the face of set-backs, and to sustain these efforts over time.

These findings provided important groundwork for the conceptual analysis developed in Chapter 4 where I detailed the social identity-based mechanisms that could underpin Disorder Identity Adaption Strategies. I argued that the fundamental dimensions of whether a condition is collectively understood by the Disorder Identity group to be positive or not (valence) and how readily the condition can be overcome (malleability) are two powerful dimensions for understanding the diversity of Disorder Identity Adaption Strategies. These ideas, in and of themselves, are not new and already some work within the social identity tradition examines these processes in a clinical context (Jetten et al., 2011). However, the distinctive contribution that I made in developing this framework was to integrate the SCT literature more thoroughly with the HPC account of mental disorder. In doing so, I showed that self-categorization processes are not merely social-psychological responses to mental disorders that in some sense pre-exist, but could actually become fundamental to the ontology of mental disorder.

More specifically, one of the key insights I made was to show that the well-established SCT process of depersonalisation, whereby group-based self-perception actively accentuates intragroup similarities on ingroup defining dimensions, could serve as a potentially powerful similarity-generating mechanism as required by HPC accounts of disorder. Moreover, the SCT account is well placed to help explain HPC dynamics too. Because the account suggests that the self-stereotyping process is context-dependent, rather

than fixed, the normative content of a Disorder Identity is predicted to systematically change in ways that maximise the contrast with outgroup representations. This context-dependent self-stereotyping could therefore serve as an externalist homeodynamic mechanism, by causing the average properties of HPC category members to systematically polarise away from outgroup representations. It thereby helps account for phenomena such as somaticizing in various contested illnesses.

I then elaborated the processes by which category members can use a shared social identity to adaptively modify their social environment. Again, the role of contested lay theories about disorder plays a pivotal role here because these theories govern how those with disorder are treated by society. The social identity account of minority social influence explains how a shared Disorder Identity provides the basis for groups to lobby for better access to treatment and resources, to change stigmatising beliefs by the outgroup, and to challenge wider social norms about the boundaries of dysfunction. However, although these identity-based social change processes have been well documented, the account of how group members systematically change themselves is somewhat less well developed. I argued that such an account needs to explain how it is that Disorder Identity groups represent stability and change to themselves and others, and how this in turn can systematically alter whether group members transition into or out of HPC disorder and whether these transitions are sustained.

Again, the SCT account provides some important insights, even if the temporal dynamics of self-categorization processes have been relatively neglected as an object of empirical research. However, in theory, the account helps illuminate the deep connections between self-categorization, social influence and attribution and how these might interact longitudinally. Specifically, the interaction between normative and comparative fit in the

self-categorization process parallels the interactive roles of covariation information and theory in contemporary accounts of attribution (Ahn et al., 1995). But the SCT analysis extends this by suggesting that attribution is often a group based process. Fellow group members provide a potentially rich resource by not only providing relevant covariation information but also a means to collectively validate and reinforce lay theories.

I therefore claimed that Disorder Identity groups often draw on lay theories to construct narratives about the group's past, present and future and that these shared narratives will predict how individuals interpret their experience, their motivations and their behaviour. Crucially, these narratives are not simply restatements of scientific accounts of HPC disorder. Indeed they may be empirically false, but nevertheless causally consequential determinants of the inflow and outflow of HPC disorder category members.

I argued that, again, the two dimensions of valence and malleability capture very important sources of variation in these group narratives and seem to be important determinants of the inflow and outflow of HPC category members. To reiterate briefly, these narratives could operate as negatively heterodynamic and negatively heterostatic HPC mechanisms to the extent that a syndrome is portrayed as a sign of positive moral virtue to be aspired to, or as a legitimate sign of human variation to be respected and perhaps even celebrated. Alternatively, narratives could also have positively heterodynamic and positively heterostatic consequences to the extent that group members perceive that a condition is a disease to be treated through biomedical intervention, an incorrigible condition that requires ongoing group support to avoid relapse, or alternatively as an opportunity to learn from, and support, fellow sufferers to ultimately overcome the symptoms of a condition. I argued that an empirical test of these claims would be to show that to the extent individuals self-categorize in terms of these Disorder Identities, then depending on the group's norms or

narratives, it is predicted that the respective HPC consequences should obtain. Of course this is qualified by the observation that for any particular condition there are usually competing, group-based narratives about how to adapt to disorder, and with varying degrees of emphasis on modifying the environment or modifying the self. Moreover, the group's norms will dynamically respond to changes in, and threats posed by, the wider social environment.

A critical test of the Disorder Identity Analysis: the collective mobility thesis

There are many possible empirical predictions that can be generated from the Disorder Identity account I have developed, but it was clearly vastly beyond the scope to examine these claims within a single thesis. Instead, I focussed on just one of the predictions that seemed to be particularly theoretically problematic. Specifically, this was the problem that it seems empirically and theoretically implausible that self-categorizing in terms of a Disorder Identity could in any sense help individuals to perceive their condition as more malleable. The very act of self-categorizing as disordered seems to implicitly reify and essentialise disorder. Indeed the link between fixedness and Disorder Identity seems to be supported by evidence within the SIT literature showing that identifying with a group is both a cause and a consequence of perceived boundary impermeability (Ellemers et al., 1997). Conversely, to the extent that the boundaries of disorder are perceived as permeable (i.e. that one's identity in this respect is malleable) then, according to SIT, members of the low status group are likely to adopt an *individual* upward mobility strategy. It would seem, then, that whatever other adaptive possibilities a Disorder Identity might enable, Disorder Identity could never provide the basis for group members to (psychologically) transition out of the group.

However, I argued that a counterpoint to this view is that there may be a relatively underexplored social identity-based process that operates in between the identity management

strategy of individual upward mobility and the clearly group-based strategies of social creativity and social competition. In the context of disorder, it is often unclear and indeed highly contested, as to whether a condition is malleable (completely permeable boundaries) or fixed (completely impermeable boundaries). Consequently, those with disorder are confronted with the questions of whether, and to what extent, change is possible. This raised the possibility that fellow disorder sufferers are perhaps uniquely placed to inform the self about the possibility of change. On this basis, I argued that a Disorder Identity may, counter-intuitively, provide the inspiration to help individuals to reduce their symptomatology and perhaps overcome disorder but that this would vary depending on whether the group held a malleable norm about the nature of the condition (the collective upward mobility thesis).

To test the collective mobility thesis I ran five studies examining five main Empirical Propositions and two Corollaries, as shown in Table 13. All of the studies examined these propositions and corollaries in the context of the anxiety disorder analogues—shyness and speech anxiety. In totality, the studies generally supported the overarching collective mobility thesis. In particular, Study 1, Study 3 and Study 5 showed broad support for Corollary 2 by demonstrating that high disorder identifiers were influenced by fellow ingroup members such that their personal beliefs about the malleability of their own condition were either more malleable or more fixed, in line with the group norm. Study 4 and Study 5 further demonstrated that the interaction between the malleability norm and disorder identification had symptomatic consequences—both in terms of reductions in self-reported shyness symptoms over a six month period, and in terms of observer-ratings of anxiety while delivering a short speech.

Notwithstanding this support for the collective mobility thesis, across the first three studies individual malleability beliefs and low group identification were consistently

correlated, in line with the individual mobility thesis. The implications of this are that fears that Disorder Identities may serve as a barrier to perceiving that a condition can change are not without foundation (Howard, 2008), but that the consequences of Disorder Identity are more complex. Specifically, the evidence presented here suggests that there may be both a main effect of disorder identification as well as a fixed norm by identification interaction effect on individual fixed beliefs. The broader clinical implications of this will be discussed shortly.

Although Studies 1-5 provided generally very consistent evidence for both collective mobility, and for individual mobility, a number of important limitations and opportunities for further research should be noted. In Study 2 I attempted to manipulate the salience of a Disorder Identity to test for the posited causal relationship between the identity by norm interaction and disorder-related behaviours and beliefs. Unfortunately, this manipulation was not successful—in large part because of the difficulties in ethically and convincingly making salient an analogue Disorder Identity. Although there was evidence for the hypothesised identification by norm interaction effect using measured identification, the current findings do not rule out the possibility that some correlate of measured identification is generating the observed effects. One possibility for avoiding this problem could be to employ a more minimal, but negatively valenced, category to demonstrate the theoretical point that upward mobility can be a collective process, rather than necessarily being individuating. Such evidence would provide a crucially important contribution to Social Identity theorising by not only elaborating the different constraints on group boundary permeability (Armenta, 2014), but by testing the boundary conditions and viability of collective upward mobility processes. This would augment the already widely researched processes of individual mobility, social creativity, and social competition.

The second major limitation of the current studies is that the use of relatively trivial disorder analogues is clearly far removed from the complex realities of clinically diagnosable mental disorder. In defence, it seemed prudent to study analogue conditions given the ethical and methodological challenges at stake. Moreover, the findings here suggest rich and fascinating possibilities for further clinical applications. I will briefly mention three of these.

Firstly, a much neglected topic within the clinical literature has been the role of vicarious modelling. As I touched on in Chapter 3, Bandura's social cognitive theory (1986) posits that, in addition to individual mastery experiences, social modelling is an important determinant of self-efficacy beliefs. However, the evidence for social modelling is weak and variable (Usher & Pajares, 2008), in large part because of uncertainties about how to operationalise the similarity of a social model. For instance, social models who share the same age and gender as an individual, do not reliably transmit self-efficacy beliefs (Grace, David, & Ryan, 2008). The collective mobility account developed here suggests a way forward by positing that it is not just any similarity that counts, what matters is whether individuals perceive that the model shares a salient social identity that is based on a theoretically relevant similarity (i.e. a shared Disorder Identity). And indeed, I tested and found support for this proposition in Study 3 by showing that people's personal malleability beliefs are influenced by the normative malleability beliefs of fellow shy ingroup members, but not by ingroup members from a theoretically unrelated group. The result is perhaps intuitively unsurprising and the idea is implicit within various interventions that attempt to socially model clinical recovery. However, group identification in interaction with norms about growth and change are seldom empirically assessed, and so the potential of this idea to support clinical change has not been fully realised.

A second application of the basic idea relates to formal group therapy programs. A much researched construct here has been group cohesiveness, which some assume should straightforwardly be linked to positive clinical outcomes. However, the evidence that group cohesiveness improves outcomes is very mixed and the cohesiveness construct is often poorly operationalised (Hornsey, Dwyer, Oei, & Dingle, 2009). Here again, the collective mobility thesis suggests that what is more important is how group members collectively negotiate and co-construct a recovery narrative and that a better way to operationalise this idea is to measure the interaction between identification with the group and the group's normative understanding about change.

The third major opportunity to extend the current studies is to examine whether such processes facilitate enduring clinical change. Although Study 4 and Study 5 provided evidence that collective mobility strategies reduce symptomatology, technically this provides evidence for a negatively homeodynamic mechanism only. What is of greater interest theoretically, and certainly clinically, is to examine whether collective mobility facilitates transition from an HPC-defined disordered state to a non-disordered state (i.e. a positively heterodynamic mechanism) and that, having identified with a Disorder Identity with a malleable norm, that this helps prevent relapse (i.e. a positively heterostatic mechanism). There is potentially much to be understood about how such transitions are collectively symbolised as a means of consolidating clinical change. It is notable that in educational contexts, for instance, leaving the group is marked symbolically with often elaborate graduation ceremonies that celebrate how the cohort has collectively changed, and emphasises how group members can leave the group as fundamentally different people (even when there is substantial individual variation in the extent of these changes). Depending on the nature of the disorder, there are interesting possibilities for examining whether such

collective ceremonies, that ritualise the process of clinical change, could help undermine, for instance, latent disorder self-schemas that are postulated to predispose individuals to relapse following successful treatment (Hedlund & Rude, 1995).

This point raises a further limitation of the current empirical studies, notably that all of the analogue conditions tested were anxiety disorders. It would be worth investigating whether different kinds of mental disorder may be amenable to these sorts of collective mobility adaption strategies in the context of clinical interventions. Anxiety disorders were chosen because of the central role that motivational processes and avoidance play in maintaining these conditions. But potentially, there may be other conditions that provide a good fit for collective mobility strategies. For instance, in the case of Major Depression, often relatively brief psychological interventions can yield substantial reductions in symptomatology, and motivational processes play an important causal role in maintaining symptoms (Abramson et al., 1981; Kneeland et al., 2016). Again, group therapy for Major Depression is not new. The novel point is that a collective mobility analysis could offer deeper insights into the mechanisms of change.

Wider implications of the Disorder Identity Analysis

The empirical studies testing the collective mobility thesis represent a small, but important first step in the broader Disorder Identity Analysis I have developed in this thesis. If nothing else, the findings illustrate how externalist HPC mechanisms can be empirically operationalised and tested. More optimistically, some might interpret the findings as yet further evidence that social groups are good for our health and wellbeing (Jetten et al., 2011). But this would miss what I hope is a much more substantive contribution from the broader analysis which is to problematise the very notion of what counts as being “good for our health and wellbeing”. Obsessively searching for mechanisms that lead to “good outcomes”

or “bad outcomes” is, in many ways, what generates the riddle of mental disorder. This is not to claim that sound empirical research cannot lead to insights that profoundly improve people’s quality of life. Clearly it can. Rather, the point is to be more cautious when extrapolating from narrow discoveries in a particular socio-historical context to making claims about timeless truths as to how best to respond to deviancy.

One example of this stems from recent debates about the negative consequences of disorder labels and whether psychologists and psychiatrists should abandon psychiatric labels altogether, because “psychiatric diagnosis increases stigma, does not aid treatment decisions, and is associated with worsening long-term prognosis” (Timimi, 2014 p. 208). Instead, several authors propose that psychiatrists and psychologists need to adopt a new way of working that promotes the recovery model, and responds to individuals with disorder as a “whole person”, rather than in terms of their diagnosis (Bracken et al., 2012; Pingani et al., 2014). What is striking in these proposals is the presumption that the relationship between stigma and labels is unambiguous, and that clinicians can straightforwardly decide to remove disorder labels and treat clients as individuals.

A recent study, examining responses from individuals who were formerly diagnosed with Aspergers Syndrome, and who are now no longer so diagnosed because the disorder has been disbanded within DSM-V, speaks to this challenge (Giles, 2014). Among a variety of responses to this decision by the DSM taskforce, many “Aspies” expressed grave fears of having their Disorder Identity taken away and being understood as neurotypical (NT). As one member from the online forum that the researchers studied put it:

Soon enough, I'll be NT. Except I won't be NT. I'll be an NT who paces constantly, touching everything in my path. I'll be an NT who has meltdowns and can't socialize.

I'll be an NT who doesn't qualify for services, has no excuse or reason for being odd, and is ... well, broken. (p. 187)

This reflection points to the serious challenges of simply assuming that disorder labels have meaning only for clinicians, and can therefore be dispensed with, and replaced with more individualised humanistic therapy. The wider implication of the Disorder Identity analysis that I have presented is that the crisis of confidence in psychiatric categories does not just play out in decisions that begin and end inside clinical consulting rooms and research laboratories. Whether clinicians and researchers like it or not, diagnostic categories seem to have become tools for creating often vibrant, sometime disturbing, and invariably complex cultures for negotiating, interacting and existing within society.

Disorder Identities will therefore continue to present both opportunities and obstacles to clinicians, and society more generally, in our attempts to control deviancy, reflect on the boundaries of normalcy, and respond sensitively to human suffering.

Table 13: *Summary of studies testing the Collective Mobility Thesis*

Disorder analogue	Study 1 Shyness	Study 2 Shyness	Study 3 Shyness	Study 4 Shyness in transition to university	Study 5 Speech anxiety
Empirical Proposition 1 (consistent with the individual mobility thesis): To the extent individuals' personal identities are made salient, they will perceive their own condition is more malleable.		Not supported. Evidence for social competition on malleability dimension under low salience conditions for low identifiers.			

Disorder analogue	Study 1 Shyness	Study 2 Shyness	Study 3 Shyness	Study 4 Shyness in transition to university	Study 5 Speech anxiety
Empirical Proposition 2 (consistent with the collective mobility thesis): To the extent individuals' Disorder Identities are salient they will be influenced by fellow disorder group members to perceive that their own condition is either relatively fixed or malleable, consistent with the group norm.		Qualified support: Identity Salience predicted social influence, but only for high identifiers.			
Corollary 1: Those who identify less strongly with a Disorder Identity group (i.e. low identifiers) will perceive their own condition to be more malleable compared to those who identify highly with the group.	Supported	Supported	Supported		

Disorder analogue	Study 1 Shyness	Study 2 Shyness	Study 3 Shyness	Study 4 Shyness in transition to university	Study 5 Speech anxiety
Corollary 2: Those who identify highly with a disorder group will be influenced by fellow disorder group members to perceive their own condition as either relatively fixed or malleable, consistent with the group norm.	Supported	Qualified support. High identifiers were socially influenced under identity salience conditions	Supported		Supported
Empirical Proposition 3: Those who identify highly with a Disorder Identity group will be influenced by fellow disorder group members endorsing malleability beliefs, but will not be similarly influenced by fellow ingroup members from theoretically irrelevant groups endorsing the same malleability beliefs about the condition.			Supported		

Disorder analogue	Study 1 Shyness	Study 2 Shyness	Study 3 Shyness	Study 4 Shyness in transition to university	Study 5 Speech anxiety
Empirical Proposition 4: To the extent an individual identifies with a disorder category, and there is a malleable group norm, then it is predicted that those individuals will display reductions in symptomatology over time.				Supported	
Empirical Proposition 5: The interaction between a malleable group norm and disorder identification will predict more malleable personal beliefs about disorder which will, in turn, positively mediate the relationship between the malleable group norm and lower symptomatology (i.e. moderated mediation).					Supported

References

- Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, 96(2), 358–372.
<http://doi.org/10.1037/0033-295X.96.2.358>
- Adams, J., & White, M. (2005). Why don't stage-based activity promotion interventions work? *Health Education Research*, 20(2), 237–243.
- Adams, S., Pill, R., & Jones, A. (1997). Medication, chronic illness and identity: The perspective of people with asthma. *Social Science & Medicine*, 45(2), 189–201.
- Agich, G. J. (1997). Toward a Pragmatic Theory of Disease. In J. M. Humber & R. F. Almeder (Eds.), *What Is Disease?* (pp. 219–246). Humana Press.
- Ahn, W., Flanagan, E. H., Marsh, J. K., & Sanislow, C. A. (2006). Beliefs About Essences and the Reality of Mental Disorders. *Psychological Science*, 17(9), 759–766.
<http://doi:10.1111/j.1467-9280.2006.01779.x>
- Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, 54(3), 299–352.
[http://doi.org/10.1016/0010-0277\(94\)00640-7](http://doi.org/10.1016/0010-0277(94)00640-7)
- Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, 33(2), 147–182.
- Akil, H., Brenner, S., Kandel, E., Kendler, K. S., King, M.-C., Scolnick, E., ... Zoghbi, H. Y. (2010). The Future of Psychiatric Research: Genomes and Neural Circuits. *Science (New York, N.Y.)*, 327(5973), 1580–1581. <http://doi.org/10.1126/science.1188654>
- Ali-Gombe, A., Guthrie, E., & McDermott, N. (1996). Mass hysteria: one syndrome or two? *The British Journal of Psychiatry*, 168(5), 633–635.

- Alloy, L. B. (1988). *Cognitive processes in depression*. Guilford Press.
- Angermeyer, M. C., & Matschinger, H. (2003). Public beliefs about schizophrenia and depression: similarities and differences. *Social psychiatry and psychiatric epidemiology*, 38(9), 526-534.
- Anthony, W. A. (1993). Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychiatric Rehabilitation Journal*, 16(4), 11-23.
- Antaki, C., Condor, S., & Levine, M. (1996). Social identities in talk: Speakers' own orientations. *British Journal of Social Psychology*, 35(4), 473-492.
- Armenta, B. (2014). Effects of social mobility perceptions on older adults' group identification. In *The psychology of social mobility for disadvantaged group members*. Amsterdam.
- Asch, S. (1952). *Social Psychology*. New York: Prentice Hall.
- Ascher, L. M., & Schotte, D. E. (1999). Paradoxical intention and recursive anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 30(2), 71-79.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21(4), 547-569.
- Bailey, J. M., Vasey, P. L., Diamond, L. M., Breedlove, S. M., Vilain, E., & Epprecht, M. (2016). Sexual orientation, controversy, and science. *Psychological Science in the Public Interest*, 17(2), 45-101.
- Bain, P. G., Hornsey, M. J., Bongiorno, R., Kashima, Y., & Crimston, D. (2013). Collective futures how projections about the future of society are related to actions and attitudes supporting social change. *Personality and Social Psychology Bulletin*, 39(4), 523-539.
- Balaratnasingam, S., & Janca, A. (2006). Mass hysteria revisited. *Current Opinion in Psychiatry*, 19(2), 171-174.

Bandura, A. (1968). A social learning interpretation of psychological dysfunctions.

Foundations of abnormal psychology, 293-344.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change.

Psychological Review, 84(2), 191.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*.

Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American*

Psychologist, 54(7), 462–479. <http://doi.org/10.1037/0003-066X.54.7.462>

Barrett, B., Muller, D., Rakel, D., Rabago, D., Marchand, L., & Scheder, J. C. (2006).

Placebo, Meaning, and Health. *Perspectives in Biology and Medicine*, 49(2), 178–

198. <http://doi.org/10.1353/pbm.2006.0019>

Barsalou, L. W. (1990). Access and inference in categorization. *Bulletin of the Psychonomic Society*, 28(3), 268-271.

Bartholomew, R. E., & Wessely, S. (2002). Protean nature of mass sociogenic illness From

possessed nuns to chemical and biological terrorism fears. *The British Journal of*

Psychiatry, 180(4), 300–306. <http://doi.org/10.1192/bjp.180.4.300>

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of

anxiety disorders: a systematic review and meta-regression. *Psychological Medicine*,

43(5), 897–910. <http://doi.org/10.1017/S003329171200147X>

Bechtel, W. (2009). Looking down, around, and up: Mechanistic explanation in psychology.

Philosophical Psychology, 22(5), 543–564.

Bechtel, W. (2013). Understanding biological mechanisms: Using illustrations from

circadian rhythm research. In *The Philosophy of Biology* (pp. 487–510). Springer.

- Bechtel, W., & Abrahamsen, A. (2008). From reduction back to higher levels. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 559–564). Retrieved from <http://csjarchive.cogsci.rpi.edu/Proceedings/2008/pdfs/p559.pdf>
- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321–333. <http://doi.org/10.1016/j.shpsa.2010.07.003>
- Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York, NY: Guilford Press; US.
- Becker, A. E., Burwell, R. A., Herzog, D. B., Hamburg, P., & Gilman, S. E. (2002). Eating behaviours and attitudes following prolonged exposure to television among ethnic Fijian adolescent girls. *The British Journal of Psychiatry*, 180(6), 509–514.
- Beebe, H., & Sabbarton-Leary, N. (2010). Are psychiatric kinds real? *European Journal of Analytic Philosophy*, 6(1), 11–27.
- Beer, J. S. (2002). Implicit Self-Theories of Shyness. *Journal of Personality & Social Psychology*, 83(4), 1009–1024.
- Beidel, D. C., & Turner, S. M. (1999). The natural course of shyness and related syndromes.
- Bell, V. (2007). Online information, extreme communities and internet therapy: Is the internet good for our mental health? *Journal of Mental Health*, 16(4), 445–457. <http://doi.org/10.1080/09638230701482378>
- Bettencourt, B., Charlton, K., Dorr, N., & Hume, D. L. (2001). Status differences and in-group bias: a meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological bulletin*, 127(4), 520.

- Beutler, L. E., Harwood, T. M., Michelson, A., Song, X., & Holman, J. (2011). Resistance/reactance level. *Journal of Clinical Psychology*, 67(2), 133-142.
- Bickle, J. (2006). Ruthless reductionism in recent neuroscience. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 36(2), 134-140.
<http://doi.org/10.1109/TSMCC.2006.871130>
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Development*, 78(1), 246-263.
<http://doi.org/10.1111/j.1467-8624.2007.00995.x>
- Blöte, A. W., Kint, M. J., Miers, A. C., & Westenberg, P. M. (2009). The relation between public speaking anxiety and social anxiety: a review. *Journal of Anxiety Disorders*, 23(3), 305-313.
- Blumner, K. H., & Marcus, S. C. (2009). Changing perceptions of depression: ten-year trends from the general social survey. *Psychiatric Services*.
- Boorse, C. (1977). Health as a Theoretical Concept. *Philosophy of Science*, 44(4), 542-573.
- Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91-121.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61(1), 127-148.
- Boyd, R. (1999). Homeostasis, species, and higher taxa. In RA Wilson (ed.), *Species: New Interdisciplinary Essays*. MIT Press, 141-85.
- Bracken, P., Thomas, P., Timimi, S., Asen, E., Behr, G., Beuster, C., ... Yeomans, D. (2012). Psychiatry beyond the current paradigm. *The British Journal of Psychiatry*, 201(6), 430-434. <http://doi.org/10.1192/bjp.bp.112.109447>

- Branscombe, N. R., Fernandez, S., Gomez, A., & Cronin, T. (2012). Moving toward or away from a group identity: Different strategies for coping with pervasive discrimination. In *The social cure: Identity, health and well-being* (pp. 115–131). New York, NY: Psychology Press; US.
- Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, 77(1), 135–149.
<http://doi.org/10.1037/0022-3514.77.1.135>
- Branthwaite, A., & Cooper, P. (1981). Analgesic effects of branding in treatment of headaches. *BMJ*, 282(6276), 1576–1578.
- Bridle, C., Riemsma, R. P., Pattenden, J., Sowden, A. J., Mather, L., Watt, I. S., & Walker, A. (2005). Systematic review of the effectiveness of health behavior interventions based on the transtheoretical model. *Psychology & Health*, 20(3), 283–301.
<http://doi.org/10.1080/08870440512331333997>
- Brown, P. M., & Turner, J. C. (2002). The role of theories in the formation of stereotype content. *Stereotypes as explanations: The formation of meaningful beliefs about social groups*, 67-89.
- Brown, R. J. (2004). Psychological mechanisms of medically unexplained symptoms: an integrative conceptual model. *Psychological Bulletin*, 130(5), 793.
- Brown, R., & Middendorf, J. (1996). The underestimated role of temporal comparison: A test of the life-span model. *The Journal of Social Psychology*, 136(3), 325.
- Brülde, B. (2007). Mental disorder and values. *Philosophy, Psychiatry, & Psychology*, 14(2), 93-102.
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64(2), 123–152.

<http://doi.org/10.1037/h0043805>

- Buckingham, S. A., Frings, D., & Albery, I. P. (2013). Group Membership and Social Identity in Addiction Recovery. *Psychology of Addictive Behaviors*, No Pagination Specified. <http://doi.org/10.1037/a0032480>
- Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., ... Ochsner, K. N. (2014). Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. *Cerebral Cortex*, 24(11), 2981–2990. <http://doi.org/10.1093/cercor/bht154>
- Burgoon, J. K., Buller, D. B., & Woodall, W. G. (Eds.). (1996). *Nonverbal communication: The unspoken dialogue*. New York: McGraw-Hill.
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655.
- Burton, P. R., Hansell, A. L., Fortier, I., Manolio, T. A., Khoury, M. J., Little, J., & Elliott, P. (2009). Size matters: just how big is BIG? Quantifying realistic sample size requirements for human genome epidemiology. *International Journal of Epidemiology*, 38(1), 263–273.
- Butler, E. A. (2011). Temporal Interpersonal Emotion Systems: The “TIES” That Form Relationships. *Personality and Social Psychology Review*, 15(4), 367–393. <http://doi.org/10.1177/1088868311411164>
- Butner, J., Diamond, L. M., & Hicks, A. M. (2007). Attachment style and two forms of affect coregulation between romantic partners. *Personal Relationships*, 14(3), 431–455. <http://doi.org/10.1111/j.1475-6811.2007.00164.x>
- Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade

of the brain: Doctrine of multilevel analysis. *American Psychologist*, 47(8), 1019.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin*, 126(6), 829–843. <http://doi.org/10.1037//0033-2909.126.6.829>

Camp, D. L., Finlay, W. M. L., & Lyons, E. (2002). Is low self-esteem an inevitable consequence of stigma? An example from women with chronic mental health problems. *Social Science & Medicine*, 55(5), 823–834. [http://doi.org/10.1016/S0277-9536\(01\)00205-2](http://doi.org/10.1016/S0277-9536(01)00205-2)

Carducci, B. J., & Golant, S. (2000). *Shyness: A bold new approach*. Harper Perennial.Chicago.

Casey, B. J., Craddock, N., Cuthbert, B. N., Hyman, S. E., Lee, F. S., & Ressler, K. J. (2013). DSM-5 and RDoC: progress in psychiatry research? *Nature Reviews Neuroscience*, 14(11), 810–814. <http://doi.org/10.1038/nrn3621>

Chabris, C. F., Lee, J. J., Benjamin, D. J., Beauchamp, J. P., Glaeser, E. L., Borst, G., ... Laibson, D. I. (2013). Why It Is Hard to Find Genes Associated With Social Science Traits: Theoretical and Empirical Considerations. *American Journal of Public Health*, 103(S1), S152–S166. <http://doi.org/10.2105/AJPH.2013.301327>

Chamberlin, J. (1978). *On our own: Patient-controlled alternatives to the mental health system*. McGraw-Hill.

Chan, S. M. (2012). Early adolescent depressive mood: Direct and indirect effects of attributional styles and coping. *Child Psychiatry & Human Development*, 43(3), 455–470.

Chanock, S. J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D. J., Thomas, G., ...

- Collins, F. S. (2007). Replicating genotype–phenotype associations. *Nature*, 447(7145), 655–660. <http://doi.org/10.1038/447655a>
- Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test Anxiety and Academic Performance in Undergraduate and Graduate Students. *Journal of Educational Psychology*, 97(2), 268–274. <http://doi.org/10.1037/0022-0663.97.2.268>
- Cheng, A. T. A. (2001). Case definition and culture: are people all the same? *The British Journal of Psychiatry*, 179(1), 1–3. <http://doi.org/10.1192/bjp.179.1.1>
- Chertoff, B. (2005) Making their skin crawl, *Popular Mechanics*, 182, 6.
- Chiu, C., Hong, Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, 73(1), 19–30. <http://doi.org/10.1037/0022-3514.73.1.19>
- Chrisler, J. C., & Caplan, P. (2002). The strange case of Dr. Jekyll and Ms. Hyde: How PMS became a cultural phenomenon and a psychiatric disorder. *Annual Review of Sex Research*, 13(1), 274–306.
- Clarke, J. N., & James, S. (2003). The radicalized self: the impact on the self of the contested nature of the diagnosis of chronic fatigue syndrome. *Social Science & Medicine*, 57(8), 1387–1395.
- Cohen, C. I. (1993). The biomedicalization of psychiatry: A critical overview. *Community Mental Health Journal*, 29(6), 509–521.
- Cooper, R. (2012). Being ill and getting better: recovery and accounts of disorder. *Recovery of People with Mental Illness: Philosophical and Related Perspectives*, 217.
- Cooper, R. (2015). Must Disorders Cause Harm? The Changing Stance of the DSM. In S. Demazeux & P. Singy (Eds.), *The DSM-5 in Perspective* (pp. 83–96). Springer

Netherlands.

- Corrigan, P. W., Kerr, A., & Knudsen, L. (2005). The stigma of mental illness: Explanatory models and methods for change. *Applied and Preventive Psychology, 11*(3), 179–190. <http://doi.org/10.1016/j.appsy.2005.07.001>
- Corrigan, P. W., Kosyluk, K. A., & Rüsch, N. (2013). Reducing self-stigma by coming out proud. *American Journal of Public Health, 103*(5), 794–800.
- Corrigan, P. W., Larson, J. E., & Ruesch, N. (2009). Self-stigma and the “why try” effect: impact on life goals and evidence-based practices. *World Psychiatry, 8*(2), 75–81.
- Corrigan, P. W., Larson, J. E., Hautamaki, J., Matthews, A., Kuwabara, S., Rafacz, J., ... & O’Shaughnessy, J. (2009). What lessons do coming out as gay men or lesbians have for people stigmatized by mental illness?. *Community mental health journal, 45*(5), 366–374.
- Corrigan, P. W., & Watson, A. C. (2002). Understanding the impact of stigma on people with mental illness. *World psychiatry, 1*(1), 16–20.
- Corrigan, P. W., Watson, A. C., & Barr, L. (2006). The Self-Stigma of Mental Illness: Implications for Self-Esteem and Self-Efficacy. *Journal of Social and Clinical Psychology, 25*(8), 875–884.
- Corstens, D., Longden, E., McCarthy-Jones, S., Waddingham, R., & Thomas, N. (2014). Emerging perspectives from the Hearing Voices Movement: implications for research and practice. *Schizophrenia bulletin, 40* (Suppl 4), S285–S294.
- Crabtree, J. W., Haslam, S. A., Postmes, T., & Haslam, C. (2010). Mental Health Support Groups, Stigma, and Self-Esteem: Positive and Negative Implications of Group Identification. *Journal of Social Issues, 66*(3), 553–569. <http://doi.org/10.1111/j.1540-4560.2010.01662.x>

- Craen, A. J. M. de, Roos, P. J., Vries, A. L. de, & Kleijnen, J. (1996). Effect of colour of drugs: systematic review of perceived effect of drugs and of their effectiveness. *BMJ*, 313(7072), 1624–1626. <http://doi.org/10.1136/bmj.313.7072.1624>
- Cramer, A. O., Waldorp, L. J., van der Maas, H. L., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, 33(2–3), 137–150.
- Craver, C. F. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, 22(5), 575–594. <http://doi.org/10.1080/09515080903238930>
- Craver, C. F. (2013). Functions and mechanisms: A perspectivalist view. In *Functions: Selection and mechanisms* (pp. 133–158). Springer Netherlands.
- Craver, C. F., & Bechtel, W. (2007). Top-down Causation Without Top-down Causes. *Biology & Philosophy*, 22(4), 547–563. <http://doi.org/10.1007/s10539-006-9028-8>
- Crazier, W. R. (1979). Shyness as a dimension of personality. *British Journal of Social and Clinical Psychology*, 18(1), 121–128. <http://doi.org/10.1111/j.2044-8260.1979.tb00314.x>
- Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, 96(4), 608.
- Cromwell, H. C., & Panksepp, J. (2011). Rethinking the cognitive revolution from a neural perspective: how overuse/misuse of the term “cognition” and the neglect of affective controls in behavioral neuroscience could be delaying progress in understanding the BrainMind. *Neuroscience & Biobehavioral Reviews*, 35(9), 2026–2035.
- Cruwys, T., Dingle, G. A., Haslam, C., Haslam, S. A., Jetten, J., & Morton, T. A. (2013). Social group memberships protect against future depression, alleviate depression symptoms and prevent depression relapse. *Social Science & Medicine*, 98, 179–186. <http://doi.org/10.1016/j.socscimed.2013.09.013>

- Cruwys, T., Dingle, G. A., Hornsey, M. J., Jetten, J., Oei, T. P. S., & Walter, Z. C. (2014). Social isolation schema responds to positive social experiences: Longitudinal evidence from vulnerable populations. *British Journal of Clinical Psychology*, 53(3), 265–280. <http://doi.org/10.1111/bjc.12042>
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, (58), 376–85.
- Cuijpers, P., van Straten, A., van Oppen, P., & Andersson, G. (2008). Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. *Journal of Clinical Psychiatry*, 69(11), 1675–1685. <http://doi.org/10.4088/JCP.v69n1102>
- Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*, 27(3), 318–326. <http://doi.org/10.1016/j.cpr.2006.11.001>
- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, 11(1), 126. <http://doi.org/10.1186/1741-7015-11-126>
- Dalrymple, K. L., & Zimmerman, M. (2013). When Does Benign Shyness Become Social Anxiety, a Treatable Disorder: The Authors Pinpoint the Fine Line between Overdiagnosis and Underdiagnosis of SAD. *Current Psychiatry*, 12(11), 21.
- David, B., & Turner, J. C. (1996). Studies in self-categorization and minority conversion: Is being a member of the out-group an advantage?. *British Journal of Social Psychology*, 35(1), 179-199.

- Davidson, L., O'Connell, M., Tondora, J., Styron, T., & Kangas, K. (2006). The top ten concerns about recovery encountered in mental health system transformation. *Psychiatric Services*, 57(5), 640-645.
- Davidson, L., & Roe, D. (2007). Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health*, 16(4), 459-470.
- Davidson, J., & Henderson, V. L. (2010). 'Coming out' on the spectrum: autism, identity and disclosure. *Social & Cultural Geography*, 11(2), 155-170.
- de Visser, R. O., Richters, J., Rissel, C., Badcock, P. B., Simpson, J. M., Smith, A. M., & Grulich, A. E. (2014). Change and stasis in sexual health and relationships: comparisons between the First and Second Australian Studies of Health and Relationships. *Sexual health*, 11(5), 505-509.
- Deacon, B. J., & Baird, G. L. (2009). The chemical imbalance explanation of depression: reducing blame at what cost?. *Journal of Social and Clinical Psychology*, 28(4), 415.
- De Castella, K., Goldin, P., Jazaieri, H., Ziv, M., Dweck, C. S., & Gross, J. J. (2013). Beliefs about emotion: Links to emotion regulation, well-being, and psychological distress. *Basic and Applied Social Psychology*, 35(6), 497-505.
- Deegan, P. E. (1990). Spirit breaking: When the helping professions hurt. *The Humanistic Psychologist*, 18(3), 301-313.
- Demoulin, S., Leyens, J.-P., & Yzerbyt, V. (2006). Lay Theories of Essentialism. *Group Processes & Intergroup Relations*, 9(1), 25-42.
<http://doi.org/10.1177/1368430206059856>
- Descartes, R. (2013). *Meditations on First Philosophy*. Broadview Press.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social

- influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3), 629–636. <http://doi.org/10.1037/h0046408>
- Diamond, L. M. (2012). The desire disorder in research on sexual orientation in women: Contributions of dynamical systems theory. *Archives of sexual behavior*, 41(1), 73–83.
- Diamond, L. M., & Rosky, C. J. (2016). Scrutinizing Immutability: Research on Sexual Orientation and US Legal Advocacy for Sexual Minorities. *The Journal of Sex Research*, 53(4-5), 363–391.
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355–391. <http://doi.org/10.1037/0033-2909.130.3.355>
- DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. Retrieved from <http://psycnet.apa.org.virtual.anu.edu.au/psycinfo/1998-06419-001>
- Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., ... others. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*, 70(1), 87–97.
- Doidge, N. (2008). *The Brain That Changes Itself: stories of personal triumph from the frontiers of brain science*. Carlton North, Vic.: Scribe.
- Doidge, N. (2015). *The Brain's Way of Healing: remarkable discoveries and recoveries from the frontiers of neuroplasticity*. Scribe.
- Dozois, D. J. A., & Westra, H. A. (2005). Development of the Anxiety Change Expectancy Scale (ACES) and validation in college, community, and clinical samples. *Behaviour Research and Therapy*, 43(12), 1655–1672. <http://doi.org/10.1016/j.brat.2004.12.001>
- Dumit, J. (2006). Illnesses you have to fight to get: Facts as forces in uncertain, emergent

- illnesses. *Social science & medicine*, 62(3), 577-590.
- Durkheim, E. (1895). *Rules of Sociological Method*. Free Press.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit Theories: Elaboration and Extension of the Model. *Psychological Inquiry*, 6(4), 322–333.
- Dweck, C. S., & Elliott-Moskwa, E. S. (2010). Self-theories: The roots of defensiveness. In *Social psychological foundations of clinical psychology* (pp. 136–153). New York, NY: Guilford Press; US.
- Dweck, C. S., Hong, Y. yi, & Chiu, C. yue. (1993). Implicit Theories Individual Differences in the Likelihood and Meaning of Dispositional Inference. *Personality and Social Psychology Bulletin*, 19(5), 644–656. <http://doi.org/10.1177/0146167293195015>
- Smith, H. C. T., & Eastham, E. J. (1973). Outbreak of abdominal pain. *The Lancet*, 302(7835), 956-958.
- Ebrahim, G. J. (1968). Mass Hysteria in School Children Notes on Three Outbreaks in East Africa. *Clinical pediatrics*, 7(7), 437-438.
- Eddy, K. T., Hennessey, M., & Thompson-Brenner, H. (2007). Eating pathology in East African women: the role of media exposure and globalization. *The Journal of Nervous and Mental Disease*, 195(3), 196–202.
- Eisenberg, D., Golberstein, E., Whitlock, J. L., & Downs, M. F. (2013). Social Contagion of Mental Health: Evidence from College Roommates. *Health Economics*, 22(8), 965–986. <http://doi.org/10.1002/hec.2873>
- Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology*, 72(3), 617.
- Ellemers, N., van Knippenberg, A., De Vries, N., & Wilke, H. (1988). Social identification

- and permeability of group boundaries. *European Journal of Social Psychology*, 18(6), 497–513. <http://doi.org/10.1002/ejsp.2420180604>
- Ellemers, N., Van Knippenberg, A., & Wilke, H. A. (1990). The influence of permeability of group boundaries and stability of group status on strategies of individual mobility and social change. *British Journal of Social Psychology*, 29(3), 233–246. <http://doi.org/http://dx.doi.org/10.1111/j.2044-8309.1990.tb00902.x>
- Elliot, A. J. (2006). The Hierarchical Model of Approach-Avoidance Motivation. *Motivation and Emotion*, 30(2), 111–116. <http://doi.org/10.1007/s11031-006-9028-7>
- Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, 196(4286), 129–136.
- Engel, G. L. (1980). The clinical application of the biopsychosocial model. *Am J Psychiatry*, 137(5), 535–544.
- Evans, G. W., Wells, N. M., & Moch, A. (2003). Housing and mental health: A review of the evidence and a methodological and conceptual critique. *Journal of Social Issues*, 59(3), 475–500.
- Farber, S. (2012). The spiritual gift of madness: The failure of psychiatry and the rise of the mad pride movement. Inner Traditions/Bear & Co.
- Farina, A., Gliha, D., Bourdreau, L. A., Ale, J. G., & Sherman, M. (1971). Mental illness and the impact of believing others know about it. *Journal of Abnormal Psychology*, 77(1), 1.
- Fernández, S., Branscombe, N. R., Gómez, Á., & Morales, J. (2012). Influence of the social context on use of surgical-lengthening and group-empowering coping strategies among people with dwarfism. *Rehabilitation Psychology*, 57(3), 224.
- Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., &

- Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature. *Psychological Medicine*, 43(3), 471–481. <http://doi.org/10.1017/S0033291712001511>
- Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, 88(3), 296–308. <http://doi.org/10.1016/j.ijpsycho.2012.10.013>
- Ferri, M., Amato, L., & Davoli, M. (1996). Alcoholics Anonymous and other 12-step programmes for alcohol dependence. In *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. Retrieved from <http://onlinelibrary.wiley.com.virtual.anu.edu.au/doi/10.1002/14651858.CD005032.pub2/abstract>
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*, 2nd. NY: McGraw-Hill, 16-15.
- Floris, J., & McPherson, S. (2015). Fighting the whole system: Dissociative identity disorder, labeling theory, and iatrogenic doubting. *Journal of Trauma & Dissociation*, 16(4), 476-493.
- Foster, J. A., & Neufeld, K.-A. M. (2013). Gut–brain axis: how the microbiome influences anxiety and depression. *Trends in Neurosciences*, 36(5), 305–312.
- Foucault, M. (1964). Folie et diraison: Histoire de la folie a l'dge classique (Paris, 1961). A much-abridged version of Histoire de lafolie was published in, 277-78.
- Fox, S., & Duggan, M. (2013). Health online 2013. *Health*, 1-55.
- Frances, A. (2013). The new crisis of confidence in psychiatric diagnosis. *Annals of Internal Medicine*, 159(3), 221–222.
- Frisch, S. (2014). How cognitive neuroscience could be more biological and what it might learn from clinical neuropsychology. *Frontiers in Human Neuroscience*, 8.

<http://doi.org/10.3389/fnhum.2014.00541>

Frost, J., & Massagli, M. (2009). PatientsLikeMe the case for a data-centered patient community and how ALS patients use the community to inform treatment decisions and manage pulmonary health. *Chronic Respiratory Disease*, 6(4), 225–229.

<http://doi.org/10.1177/1479972309348655>

Fulford, K. W., & Sadler, J. Z. (2000). History and philosophy. *Current Opinion in Psychiatry*, 13(6), 679–681.

Gallagher, A., Bridgeman, B., & Cahalan, C. (2002). The Effect of Computer-Based Tests on Racial-Ethnic and Gender Groups. *Journal of Educational Measurement*, 39(2), 133–147. <http://doi.org/10.1111/j.1745-3984.2002.tb01139.x>

Gartrell, N. K., Bos, H. M., & Goldberg, N. G. (2012). New trends in same-sex sexual contact for American adolescents?. *Archives of Sexual Behavior*, 41(1), 5-7.

Gelman, S. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8(9), 404–409. <http://doi.org/10.1016/j.tics.2004.07.001>

Gelman, S. A. (2009). LEARNING FROM OTHERS: CHILDREN'S CONSTRUCTION OF CONCEPTS. *Annual Review of Psychology*, 60, 115–140. <http://doi.org/10.1146/annurev.psych.59.103006.093659>

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23(3), 183–209. [http://doi.org/10.1016/0010-0277\(86\)90034-X](http://doi.org/10.1016/0010-0277(86)90034-X)

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, 38(3), 213–244. [http://doi.org/10.1016/0010-0277\(91\)90007-](http://doi.org/10.1016/0010-0277(91)90007-)

Q

Gerard, R. W. (1955). The academic lecture: The biological roots of psychiatry. *American Journal of Psychiatry*, 112(2), 81-90.

- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American psychologist*, 40(3), 266.
- Ghaemi, S. N. (2009). The rise and fall of the biopsychosocial model. *The British Journal of Psychiatry*, 195(1), 3–4. <http://doi.org/10.1192/bjp.bp.109.063859>
- Giles, D. C. (2014). “DSM-V is taking away our identity”: The reaction of the online community to the proposed changes in the diagnosis of Asperger’s disorder. *Health*, 18(2), 179–195. <http://doi.org/10.1177/1363459313488006>
- Giosan, C., Glovsky, V., & Haslam, N. (2001). The lay concept of ‘mental disorder’: a cross-cultural study. *Transcultural Psychiatry*, 38(3), 317-332.
- Glovsky, V., & Haslam, N. (2003). Acculturation and changing concepts of mental disorder: Brazilians in the USA. *Transcultural psychiatry*, 40(1), 50-61.
- Godman, M. (2013). Psychiatric Disorders qua Natural Kinds: The Case of the “Apathetic Children.” *Biological Theory*, 7(2), 144–152. <http://doi.org/10.1007/s13752-012-0057-z>
- Goffman, E. (1963). Notes on the Management of Spoiled Identity. *New York: Simon & Shuster. Goffman Stigma: Notes on the Management of Spoiled Identity 1963.*
- Gove, W. R. (1975). The Labelling Theory of Mental Illness: A Reply to Scheff. *American Sociological Review*, 40(2), 242. <http://doi.org/10.2307/2094348>
- Grace, D. M., David, B. J., & Ryan, M. K. (2008). Investigating Preschoolers’ Categorical Thinking About Gender Through Imitation, Attention, and the Use of Self-Categories. *Child development*, 79(6), 1928-1941.
- Grebb, J. A., & Carlsson, A. (2009). Introduction and considerations for a brain-based diagnostic system in psychiatry. In Kaplan & Sadock’s *Comprehensive Textbook of Psychiatry*, 9, 1-4.

- Gutheil, G., Gelman, S. A., Klein, E., Michos, K., & Kelaita, K. (2008). Preschoolers' use of spatiotemporal history, appearance, and proper name in determining individual identity. *Cognition*, 107(1), 366–380. <http://doi.org/10.1016/j.cognition.2007.07.014>
- Guze, S. B. (1989). Biological psychiatry: is there any other kind? *Psychological Medicine*, 19(2), 315–323.
- Hacking, I. (1998). *Mad travelers: Reflections on the reality of transient mental illnesses*. University of Virginia Press.
- Hacking, I. (1999). *The social construction of what?*. Harvard university press.
- Hahn, T., Kircher, T., Straube, B., Wittchen, H.-U., Konrad, C., Ströhle, A., ... others. (2015). Predicting Treatment Response to Cognitive Behavioral Therapy in Panic Disorder With Agoraphobia by Integrating Local Neural Information. *JAMA Psychiatry*, 72(1), 68–74.
- Hall, W., Carter, A., & Forlini, C. (2015). The brain disease model of addiction: is it supported by the evidence and has it delivered on its promises?. *The Lancet Psychiatry*, 2(1), 105-110.
- Harkins, S. G. (2006). Mere effort as the mediator of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, 91(3), 436–455. <http://doi.org/10.1037/0022-3514.91.3.436>
- Harris, M. J., Milich, R., Corbitt, E. M., Hoover, D. W., & Brady, M. (1992). Self-fulfilling effects of stigmatizing information on children's social interactions. *Journal of Personality and Social Psychology*, 63(1), 41.
- Harris, W. (2013). *Mental Disorders in the Classical World*. Boston: BRILL.
- Harrison, G., Hopper, K., Craig, T., Laska, E., Siegel, C., Wanderling, J., ... others. (2001). Recovery from psychotic illness: a 15-and 25-year international follow-up study. *The*

British Journal of Psychiatry, 178(6), 506–517.

Haslam, N. (2002a). Essentialist thinking about depression: Evidence for polarized beliefs.

[References]. *Psychological Reports*, 1253–1254.

<http://doi.org/http://dx.doi.org/10.2466/PRO.91.8.1253-1254>

Haslam, N. (2002b). Practical, Functional, and Natural Kinds: Comment. *Philosophy*,

Psychiatry, & Psychology, 9(3), 237–241.

Haslam, N. (2005). Dimensions of Folk Psychiatry. *Review of General Psychology*, 9(1),

35–47. <http://doi.org/10.1037/1089-2680.9.1.35>

Haslam, N. (2009). The taxon concept is not taxonic: Response to Grove (2008).

Psychological Reports, 104(3), 784–786.

Haslam, N. (2011). Genetic Essentialism, Neuroessentialism, and Stigma: Commentary on

Dar-Nimrod and Heine (2011). *Psychological Bulletin*, 137(5), 819–824.

Haslam, N. (2014). 2 Natural Kinds in Psychiatry: Conceptually Implausible, Empirically

Questionable, and Stigmatizing. *Classifying Psychopathology: Mental Kinds and Natural Kinds*, 11.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology.

Psychological Inquiry, 27(1), 1–17.

Haslam, N., Bastian, B., Bain, P., & Kashima, Y. (2006). Psychological Essentialism,

Implicit Theories, and Intergroup Relations. *Group Processes & Intergroup Relations*, 9(1), 63–76. <http://doi.org/10.1177/1368430206059861>

Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist Beliefs about Personality and

Their Implications. *Personality and Social Psychology Bulletin*, 30(12), 1661–1673. <http://doi.org/10.1177/0146167204271182>

Haslam, N., & Ernst, D. (2002). Essentialist beliefs about mental disorders. *Journal of*

Social and Clinical Psychology, 21(6), 628–644.

Haslam, N., & Kvaale, E. P. (2015). Biogenetic Explanations of Mental Disorder The Mixed-Blessings Model. *Current Directions in Psychological Science*, 24(5), 399–404.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, 39(1), 113–127.

Haslam, S. A., Jetten, J., O'Brien, A., & Jacobs, E. (2004). Social identity, social influence and reactions to potentially stressful tasks: support for the self-categorization model of stress. *Stress and Health*, 20(1), 3–9. <http://doi.org/10.1002/smi.995>

Haslam, S. A., Jetten, J., & Waghorn, C. (2009). Social identification, stress and citizenship in teams: a five-phase longitudinal study. *Stress and Health*, 25(1), 21–30. <http://doi.org/10.1002/smi.1221>

Haslam, S., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social identity salience and the emergence of stereotype consensus. *Personality and Social Psychology Bulletin*, 25(7), 809–818. <http://doi.org/http://dx.doi.org/10.1177/0146167299025007004>

Haslam, S., Oakes, P. J., Turner, J. C., & McGarty, C. (1995). Social categorization and group homogeneity: Changes in the perceived applicability of stereotype content as a function of comparative context and trait favourableness. *British Journal of Social Psychology*, 34(2), 139–160. <http://doi.org/http://dx.doi.org/10.1111/j.2044-8309.1995.tb01054.x>

Haslam, S. A., O'Brien, A., Jetten, J., Vormedal, K., & Penna, S. (2005). Taking the strain: Social identity, social support, and the experience of stress. *British Journal of Social Psychology*, 44(3), 355–370. <http://doi.org/10.1348/014466605X37468>

Haslam, S. A., Reicher, S. D., & Levine, M. (2011). When other people are heaven, when

other people are hell: How social identity determines the nature and impact of social support. *The Social Cure: Identity, Health, and Well-Being*, 157–174.

Haslam, S. A., Reicher, S. D., & Platow, M. J. (2010). *The new psychology of leadership: Identity, influence and power*. Psychology Press.

Haslam, S. A., Turner, J. C., Oakes, P. J., Reynolds, K. J., & Doosje, B. (2002). From personal pictures in the head to collective tools in the world: How shared stereotypes allow groups to represent and change social reality. Retrieved from <http://psycnet.apa.org/psycinfo/2003-02633-008>

Haslam, S., & Turner, J. C. (1992). Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation. *European Journal of Social Psychology*, 22(3), 251–277. <http://doi.org/http://dx.doi.org/10.1002/ejsp.2420220305>

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. *Manuscript Submitted for Publication*. Retrieved from <http://www.personal.psu.edu/jxb14/M554/articles/process2012.pdf>

Hedlund, S., & Rude, S. S. (1995). Evidence of latent depressive schemas in formerly depressed individuals. *Journal of Abnormal Psychology*, 104(3), 517–525. <http://doi.org/10.1037/0021-843X.104.3.517>

Hein, R., Beckmann, L., & Chang-Claude, J. (2008). Sample size requirements for indirect association studies of gene–environment interactions ($G \times E$). *Genetic Epidemiology*, 32(3), 235–245. <http://doi.org/10.1002/gepi.20298>

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, 12(4), 748–762.

<http://doi.org/10.1037/a0025036>

Henderson, L., & Zimbardo, P. G. (1998). An overview of Shyness, what is known and how it is treated.

Henderson, L., & Zimbardo, P. G. (2001). *Shyness as a clinical condition: The Stanford model*.

Henderson, L., Zimbardo, P., & Carducci, B. (2010). Shyness. *Corsini Encyclopedia of Psychology*. Chicago.

Henry, J. D., von Hippel, C., & Shapiro, L. (2010). Stereotype threat contributes to social difficulties in people with schizophrenia. *British Journal of Clinical Psychology*, 49(1), 31–41. <http://doi.org/http://dx.doi.org/10.1348/014466509X421963>

Hester, R. K., Delaney, H. D., & Campbell, W. (2011). Moderate Drinking and moderation management: Outcomes of a randomized clinical trial with non-dependent problem drinkers. *Journal of Consulting and Clinical Psychology*, 79(2), 215.

Hibel, L. C., Granger, D. A., Blair, C., & Cox, M. J. (2009). Intimate partner violence moderates the association between mother–infant adrenocortical activity across an emotional challenge. *Journal of Family Psychology*, 23(5), 615–625.
<http://doi.org/10.1037/a0016323>

Higgins, R. L., Berglas, S., & Snyder, C. R. (1990). *Self-handicapping: The paradox that isn't*.

Hodgins, D. (2005). Can Patients with Alcohol Use Disorders Return to Social Drinking? Yes, So What Should We Do about It? *Canadian Journal of Psychiatry*, 50(5), 264.

Hogan, M. F. (2003). New Freedom Commission report: the President's New Freedom Commission: recommendations to transform mental health care in America. *Psychiatric Services*, 54(11), 1467–1474.

- Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European review of social psychology*, 11(1), 223-255.
- Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication theory*, 16(1), 7-30.
- Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, 255-269.
- Hogg, M. A., Turner, J. C., & Davidson, B. (1990). Polarized Norms and Social Frames of Reference: A Test of the Self-Categorization Theory of Group Polarization. *Basic and Applied Social Psychology*, 11(1), 77-100.
http://doi.org/10.1207/s15324834baspl101_6
- Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review*, 65(2), 117-127.
- Holmes, L. (2014). 6 Things Shy People Can Teach Us About Success. *Huffington Post*.
- Holt-Lunstad, J., & Smith, T. B. (2012). Social Relationships and Mortality. *Social and Personality Psychology Compass*, 6(1), 41-53. <http://doi.org/10.1111/j.1751-9004.2011.00406.x>
- Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit Theories, Attributions, and Coping: A Meaning System Approach. *Journal of Personality*, 77(3), 588-599.
- Hopko, D. R., Stowell, J., Jones, W. H., Armento, M. E. A., & Cheek, J. M. (2005). Psychometric Properties of the Revised Cheek and Buss Shyness Scale. *Journal of Personality Assessment*, 84(2), 185-192.

http://doi.org/10.1207/s15327752jpa8402_08

- Hornsey, M. J., Dwyer, L., Oei, T. P., & Dingle, G. A. (2009). Group processes and outcomes in group psychotherapy: Is it time to let go of “cohesiveness”? *International Journal of Group Psychotherapy*, 59(2), 267–278. <http://doi.org/Comment/Reply>
- Houts, A. C. (2001). Harmful dysfunction and the search for value neutrality in the definition of mental disorder: response to Wakefield, part 2. *Behaviour Research and Therapy*, 39(9), 1099–1132. [http://doi.org/10.1016/S0005-7967\(01\)00053-5](http://doi.org/10.1016/S0005-7967(01)00053-5)
- Howard, J. (2006). Expecting and Accepting: The Temporal Ambiguity of Recovery Identities. *Social Psychology Quarterly*, 69(4), 307–324. <http://doi.org/10.1177/019027250606900402>
- Howard, J. (2008). Negotiating an Exit: Existential, Interactional, and Cultural Obstacles to Disorder Disidentification. *Social Psychology Quarterly*, 71(2), 177–192.
- Hucklenbroich, P. (2014). Medical criteria of pathologicity and their role in scientific psychiatry—comments on the articles of Henrik Walter and Marco Stier. *Frontiers in Psychology*, 5. <http://doi.org/10.3389/fpsyg.2014.00128>
- Hume, D. (2012 / 1739). *A treatise of human nature*. Courier Corporation.
- Hunter, J. A., Platow, M. J., Bell, L. M., Kypri, K., & Lewis, C. A. (1997). Intergroup bias and self-evaluation: Domain-specific self-esteem, threats to identity and dimensional importance. *British Journal of Social Psychology*, 36(4), 405–426. <http://doi.org/10.1111/j.2044-8309.1997.tb01141.x>
- Hutto, B. (2001). Syphilis in Clinical Psychiatry: A Review. *Psychosomatics*, 42(6), 453–460. <http://doi.org/10.1176/appi.psy.42.6.453>
- Iacono, W. G., Vaidyanathan, U., Vrieze, S. I., & Malone, S. M. (2014). Knowns and unknowns for psychophysiological endophenotypes: Integration and response to

- commentaries: Endophenotype knowns and unknowns. *Psychophysiology*, 51(12), 1339–1347. <http://doi.org/10.1111/psyp.12358>
- Ickes, W., & Layden, M. A. (1978). Attributional styles. *New directions in attribution research*, 2, 119–152.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167(7), 748–751.
- Insel, T. R., & Scolnick, E. M. (2006). Cure therapeutics and strategic prevention: raising the bar for mental health research. *Molecular Psychiatry*, 11(1), 11–17. <http://doi.org/10.1038/sj.mp.4001777>
- Insel TR, & Wang PS. (2010). Rethinking mental illness. *JAMA*, 303(19), 1970–1971. <http://doi.org/10.1001/jama.2010.555>
- Ioannidis, J. P. A. (2007). Non-Replication and Inconsistency in the Genome-Wide Association Setting. *Human Heredity*, 64(4), 203–213. <http://doi.org/10.1159/000103512>
- Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of personality and social psychology*, 50(6), 1141.
- Jaarsma, P., & Welin, S. (2011). Autism as a Natural Human Variation: Reflections on the Claims of the Neurodiversity Movement. *Health Care Analysis*, 20(1), 20–30. <http://doi.org/10.1007/s10728-011-0169-9>
- Jablensky, A., & Waters, F. (2014). RDoC: a roadmap to pathogenesis? *World Psychiatry*, 13(1), 43–44. <http://doi.org/10.1002/wps.20100>
- Jackson, L. A., Sullivan, L. A., Harnish, R., & Hodge, C. N. (1996). Achieving positive social identity: Social mobility, social creativity, and permeability of group

boundaries. *Journal of Personality and Social Psychology*, 70(2), 241–254.

<http://doi.org/10.1037/0022-3514.70.2.241>

Jacobson, N., & Greenley, D. (2001). What is recovery? A conceptual model and explication. *Psychiatric Services*, 52(4), 482–485.

Jamieson, J. P., Nock, M. K., & Mendes, W. B. (2012). Mind over matter: reappraising arousal improves cardiovascular and cognitive responses to stress. *Journal of Experimental Psychology: General*, 141(3), 417.

Jeste, D. V., del Carmen, R., Lohr, J. B., & Wyatt, R. J. (1985). Did schizophrenia exist before the eighteenth century? *Comprehensive Psychiatry*, 26(6), 493–503.

[http://doi.org/10.1016/0010-440X\(85\)90016-1](http://doi.org/10.1016/0010-440X(85)90016-1)

Jetten, J., Haslam, C., & Haslam, S. A. (2011). *The Social Cure: Identity, Health and Well-Being*. Hove: Psychology Press.

Jetten, J., Iyer, A., Tsivrikos, D., & Young, B. M. (2008). When is individual mobility costly? The role of economic and social identity factors. *European Journal of Social Psychology*, 38(5), 866–879. <http://doi.org/10.1002/ejsp.471>

Joiner, T. E., & Katz, J. (1999). Contagion of Depressive Symptoms and Mood: Meta-analytic Review and Explanations From Cognitive, Behavioral, and Interpersonal Viewpoints. *Clinical Psychology: Science and Practice*, 6(2), 149–164.

<http://doi.org/10.1093/clipsy.6.2.149>

Johnson, A. M., Mercer, C. H., Erens, B., Copas, A. J., McManus, S., Wellings, K., ... & Purdon, S. (2001). Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. *The Lancet*, 358(9296), 1835–1842.

Jones, T. F., Craig, A. S., Hoy, D., Gunter, E. W., Ashley, D. L., Barr, D. B., ... & Schaffner, W. (2000). Mass psychogenic illness attributed to toxic exposure at a high school.

New England Journal of Medicine, 342(2), 96-100.

Joseph, J. (2012). The “Missing Heritability” of Psychiatric Disorders: Elusive Genes or Non-Existent Genes? *Applied Developmental Science*, 16(2), 65–83.
<http://doi.org/10.1080/10888691.2012.667343>

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, 25(6), 881–919.

Kalish, C. (1998). Natural and artifactual kinds: Are children realists or relativists about categories? *Developmental Psychology*, 34(2), 376–391. <http://doi.org/10.1037/0012-1649.34.2.376>

Kandel, E. R. (1998). A new intellectual framework for psychiatry. *American Journal of Psychiatry*, 155(4), 457–469.

Kappes, A., & Schikowski, A. (2013). Implicit theories of emotion shape regulation of negative affect. *Cognition & Emotion*, 27 (5), 952–960.

Kasper, S., & Dold, M. (2015). Factors contributing to the increasing placebo response in antidepressant trials. *World Psychiatry*, 14(3), 304-306.

Katschnig, H. (2010). Are psychiatrists an endangered species? Observations on internal and external challenges to the profession. *World Psychiatry*, 9(1), 21-28.

Kelley, H. (1967). Attribution Theory in Social Interaction (Vol. 15).

Kellogg, S. (1993). Identity and recovery. *Psychotherapy: Theory, Research, Practice, Training*, 30(2), 235.

Kemp, J. J., Lickel, J. J., & Deacon, B. J. (2014). Effects of a chemical imbalance causal explanation on individuals' perceptions of their depressive symptoms. *Behaviour Research and Therapy*, 56, 47-52.

- Kendell, R., & Jablensky, A. (2003). Distinguishing Between the Validity and Utility of Psychiatric Diagnoses. *American Journal of Psychiatry*, 160(1), 4–12.
<http://doi.org/10.1176/appi.ajp.160.1.4>
- Kendler, K. S. (2005a). “A Gene for...”: The Nature of Gene Action in Psychiatric Disorders. *American Journal of Psychiatry*, 162(7), 1243–1252.
<http://doi.org/10.1176/appi.ajp.162.7.1243>
- Kendler, K. S. (2005b). Toward a Philosophical Structure for Psychiatry. *The American Journal of Psychiatry*, 162(3), 433–40.
- Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine*, 41(6), 1143–1150.
<http://doi.org/10.1017/S0033291710001844>
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. *The Handbook of Social Psychology*, 1(4), 233–265.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, 62(6), 593-602.
- Khan, A., & Brown, W. A. (2015). Antidepressants versus placebo in major depression: an overview. *World Psychiatry*, 14(3), 294-300.
- Khan, S. S., Hopkins, N., Reicher, S., Tewari, S., Srinivasan, N., & Stevenson, C. (2014). Shared identity predicts enhanced health at a mass gathering. *Group Processes & Intergroup Relations*, 1368430214556703. <http://doi.org/10.1177/1368430214556703>
- Kinderman, P. (2014) *A Prescription for Psychiatry. Why We Need a Whole New Approach to Mental Health and Wellbeing*. Basingstoke, UK: Palgrave MacMillan.

- King, M. (2003). Dropping the diagnosis of homosexuality: did it change the lot of gays and lesbians in Britain?. *Australian and New Zealand Journal of Psychiatry*, 37(6), 684-688.
- Kingdon, D., & Young, A. H. (2007). Research into putative biological mechanisms of mental disorders has been of no value to clinical psychiatry. *The British Journal of Psychiatry*, 191(4), 285–290. <http://doi.org/10.1192/bjp.bp.107.035675>
- Kinnier Wilson, J. V. (1967). Mental diseases of ancient Mesopotamia. *Diseases in Antiquity: A Survey of the Diseases, Injuries and Surgery of Early Populations*. Springfield: Charles C. Thomas, 723–33.
- Kirk, S. A., Gomory, T., & Cohen, D. (2013). *Mad Science: Psychiatric Coercion, Diagnosis, and Drugs*. Transaction Publishers.
- Kirsch, I. (2014). The Emperor's New Drugs: Medication and Placebo in the Treatment of Depression. In F. Benedetti, P. Enck, E. Frisaldi, & M. Schedlowski (Eds.), *Placebo* (pp. 291–303). Springer Berlin Heidelberg. Retrieved from http://link.springer.com.virtual.anu.edu.au/chapter/10.1007/978-3-662-44519-8_16
- Kirsch, I., Moore, T. J., Scoboria, A., & Nicholls, S. S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration. *Prevention & Treatment*, 5(1). <http://doi.org/10.1037/1522-3736.5.1.523a>
- Kit, K. A., Tuokko, H. A., & Mateer, C. A. (2008). A Review of the Stereotype Threat Literature and Its Application in a Neurological Population. *Neuropsychology Review*, 18(2), 132–148. <http://doi.org/10.1007/s11065-008-9059-9>
- Klauer, T., Ferring, D., & Filipp, S. H. (1998). “Still stable after all This...?”: temporal comparison in coping with severe and chronic disease. *International Journal of*

Behavioral Development, 22(2), 339-355.

Klorman, R., Hilpert, P. L., Michael, R., LaGana, C., & Sveen, O. B. (1980). Effects of coping and mastery modeling on experienced and inexperienced pedodontic patients' disruptiveness. *Behavior Therapy*, 11(2), 156-168.

Kneeland, E. T., Dovidio, J. F., Joormann, J., & Clark, M. S. (2016a). Emotion malleability beliefs, emotion regulation, and psychopathology: Integrating affective and clinical science. *Clinical psychology review*, 45, 81-88.

Kneeland, E. T., Nolen-Hoeksema, S., Dovidio, J. F., & Gruber, J. (2016b). Emotion malleability beliefs influence the spontaneous regulation of social anxiety. *Cognitive Therapy and Research*, 1-14.

Kohut, H. (2009). *The Restoration of the Self*. University of Chicago Press.

Koblenzer, C. S. (2006). The challenge of Morgellons disease. *Journal of the American Academy of Dermatology*, 55(5), 920-922.

Kornhaber, R. C., & Schroeder, H. E. (1975). Importance of model similarity on extinction of avoidance behavior in children. *Journal of Consulting and Clinical Psychology*, 43(5), 601.

Kovács, J. (1989). Concepts of Health and Disease. *Journal of Medicine and Philosophy*, 14(3), 261-267. <http://doi.org/10.1093/jmp/14.3.261>

Kubler-Ross, E., & Kessler, D. (2009). The five stages of grief. Retrieved from *World Wide Web* November, 3, 2009.

Kuorikoski, J., & Pöyhönen, S. (2012). Looping kinds and social mechanisms. *Sociological Theory*, 30(3), 187-205.

Kurosawa, K., & Harackiewicz, J. M. (1995). Test Anxiety, Self-Awareness, and Cognitive Interference: A Process Analysis. *Journal of Personality*, 63(4), 931-951.

<http://doi.org/10.1111/j.1467-6494.1995.tb00321.x>

Kvaale, E. P., Gottdiener, W. H., & Haslam, N. (2013). Biogenetic explanations and stigma: A meta-analytic review of associations among laypeople. *Social Science & Medicine*, 96, 95-103.

Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013). The “side effects” of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review*, 33(6), 782–794. <http://doi.org/10.1016/j.cpr.2013.06.002>

Kwan, S. (2009). Framing the Fat Body: Contested Meanings between Government, Activists, and Industry*. *Sociological Inquiry*, 79(1), 25–50.
<http://doi.org/10.1111/j.1475-682X.2008.00271.x>

Laing, R. D., Esterson, A., & Cooper, D. G. (1970). *Sanity, madness, and the family: Families of schizophrenics*. Penguin.

Lakhan, S. E., Vieira, K., & Hamlat, E. (2010). Biomarkers in psychiatry: drawbacks and potential for misuse. *International Archives of Medicine*, 3(1), 1.
<http://doi.org/10.1186/1755-7682-3-1>

Laland, K. N., & Sterelny, K. (2006). Perspective: Seven Reasons (not) to Neglect Niche Construction. *Evolution*, 60(9), 1751–1762. <http://doi.org/10.1111/j.0014-3820.2006.tb00520.x>

Lane, C. (2006). How Shyness Became an Illness: A Brief History of Social Phobia. *Common Knowledge*, 12(3), 388–409.

Lauren B. Alloy, L. Y. A. (1988). The Hopelessness Theory of Depression: Attributional Aspects. *The British Journal of Clinical Psychology / the British Psychological Society*, 27 (Pt 1), 5–21. <http://doi.org/10.1111/j.2044-8260.1988.tb00749.x>

Lazarus, R. S., & Folkman, S. (1984). Stress. *Appraisal, and Coping*, 456.

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L., Pennekamp, S. F., Doosje, B., ...

Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, 95(1), 144–165. <http://doi.org/http://dx.doi.org/10.1037/0022-3514.95.1.144>

Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice*, 21(1), 67-83.

Lebowitz, M. S., Pyun, J. J., & Ahn, W. K. (2014). Biological explanations of generalized anxiety disorder: effects on beliefs about prognosis and responsibility. *Psychiatric Services*, 65(4), 498-503.

Lemaine, G. (1974). Social differentiation and social originality. *European Journal of Social Psychology*, 4(1), 17–52. <http://doi.org/10.1002/ejsp.2420040103>

Lembke, A., & Humphreys, K. (2012). Moderation Management: A Mutual-Help Organization for Problem Drinkers Who Are Not Alcohol-Dependent. *Journal of Groups in Addiction & Recovery*, 7(2–4), 130–141. <http://doi.org/10.1080/1556035X.2012.705657>

Lennon, M. C., Link, B. G., Marbach, J. J., & Dohrenwend, B. P. (1989). The stigma of chronic facial pain and its impact on social relationships. *Social Problems*, 36(2), 117–134.

Levine, R. M. (1999). Identity and illness: The effects of identity salience and frame of reference on evaluation of illness and injury. *British Journal of Health Psychology*, 4(1), 63-80.

Levine, R. M., & Reicher, S. D. (1996). Making sense of symptoms: Self-categorization and

- the meaning of illness and injury. *British Journal of Social Psychology*, 35(2), 245-256.
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6), 1421.
- Levy, S. R., Plaks, J. E., Hong, Y. Y., Chiu, C. Y., & Dweck, C. S. (2001). Static versus dynamic theories and the perception of groups: Different routes to different destinations. *Personality and Social Psychology Review*, 5(2), 156-168.
- Leyens, J.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M.-P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, 31(4), 395-411. <http://doi.org/10.1002/ejsp.50>
- Lieberman, R. P., & Kopelowicz, A. (2005). Recovery from schizophrenia: a concept in search of research. *Psychiatric Services*.
- Lilienfeld, S. O. (n.d.). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy*. <http://doi.org/10.1016/j.brat.2014.07.019>
- Link, B. G. (1987). Understanding labeling effects in the area of mental disorders: An assessment of the effects of expectations of rejection. *American Sociological Review*, 52(1), 96-112. <http://doi.org/http://dx.doi.org/10.2307/2095395>
- Link, B. G., Cullen, F. T., Struening, E., Shrout, P. E., & Dohrenwend, B. P. (1989). A Modified Labeling Theory Approach to Mental Disorders: An Empirical Assessment. *American Sociological Review*, 54(3), 400-423.
- Link, B. G., Mirotznik, J., & Cullen, F. T. (1991). The effectiveness of stigma coping

orientations: Can negative consequences of mental illness labeling be avoided?

Journal of Health and Social Behavior, 302–320.

Link, B. G., & Phelan, J. C. (2006). Stigma and its public health implications. *The Lancet*, 367(9509), 528–529.

Link, B. G., Struening, E. L., Rahav, M., Phelan, J. C., & Nuttbrock, L. (1997). On Stigma and Its Consequences: Evidence from a Longitudinal Study of Men with Dual Diagnoses of Mental Illness and Substance Abuse. *Journal of Health and Social Behavior*, 38(2), 177. <http://doi.org/10.2307/2955424>

Liu, J. H., & László, J. (2007). A Narrative Theory of History and Identity. In *Social Representations and Identity* (pp. 85-107). Palgrave Macmillan US.

Livingston, J. D., & Boyd, J. E. (2010). Correlates and consequences of internalized stigma for people living with mental illness: A systematic review and meta-analysis. *Social Science & Medicine*, 71(12), 2150–2161. <http://doi.org/10.1016/j.socscimed.2010.09.030>

Lupton, D. (2013). *Fat politics: Collected writings*. Available at SSRN 2273419.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, 3(1), 23–48.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33(2), 101-121.

Malt, B. C. (1990). Features and beliefs in the mental representation of categories. *Journal of Memory and Language*, 29(3), 289–315. [http://doi.org/10.1016/0749-596X\(90\)90002-H](http://doi.org/10.1016/0749-596X(90)90002-H)

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., ... others. (2009). Finding the missing heritability of complex diseases. *Nature*,

461(7265), 747–753.

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3), 301–306. <http://doi.org/10.1111/j.2044-8260.1992.tb00997.x>

Masland, W. S. (2006). Letter. *Psychiatric services*, 57(10), 1510-1510.

McCroskey, J. C. (1978). Validity of the PRCA as an index of oral communication apprehension. *Communication Monographs*, 45(3), 192–203.
<http://doi.org/10.1080/03637757809375965>

McFall, S. R., Jamieson, J. P., & Harkins, S. G. (2009). Testing the mere effort account of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, 96(1), 135–154. <http://doi.org/http://dx.doi.org/10.1037/a0012878>

McGarty, C. (1999). *Categorization in social psychology*. Sage.

McGarty, C. (2001). Social Identity Theory does not maintain that identification produces bias, and Self-categorization Theory does not maintain that salience is identification: Two comments on Mummendey, Klink and Brown. *British Journal of Social Psychology*, 40(2), 173–176.

McGarty, C., & Grace, D. M. (1999). The constraints of the social context on categorization. In C. McGarty (Ed.), *Categorization in social psychology* (pp. 190–209). London: Sage.

McGarty, C., Yzerbyt, V. Y., & Spears, R. (2002). *Stereotypes as explanations: The formation of meaningful beliefs about social groups*. Cambridge University Press.

McLaren, N. (2006). Interactive dualism as a partial solution to the mind–brain problem for psychiatry. *Medical Hypotheses*, 66(6), 1165–1173.

<http://doi.org/10.1016/j.mehy.2005.12.023>

McNally, R. J. (2012). The Ontology of Posttraumatic Stress Disorder: Natural Kind, Social Construction, or Causal System? *Clinical Psychology: Science and Practice*, 19(3), 220–228. <http://doi.org/10.1111/cpsp.12001>

McWade, B., Milton, D., & Beresford, P. (2015). Mad studies and neurodiversity: a dialogue. *Disability & Society*, 30(2), 305–309.

Mead, G. H. (1932). *Mind Self and Society from the Standpoint of a Social Behaviorist*. University of Chicago, Chicago.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York, NY, US: Cambridge University Press.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist*, 50(4), 266–275.
<http://doi.org/10.1037/0003-066X.50.4.266>

Meehl, P. E., & Yonce, L. J. (1994). Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports*. Retrieved from
<http://psycnet.apa.org.virtual.anu.edu.au/psycinfo/1995-08099-001>

Meehl, P. E., & Yonce, L. J. (1996). Taxometric analysis: II. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator (MAXCOV procedure). *Psychological Reports*, 78(3c), 1091–1227.

Mitte, K. (2005). Meta-Analysis of Cognitive-Behavioral Treatments for Generalized Anxiety Disorder: A Comparison With Pharmacotherapy. *Psychological Bulletin*, 131(5), 785–795. <http://doi.org/10.1037/0033-2909.131.5.785>

- Molden, D. C., & Dweck, C. S. (2006). Finding “Meaning” in Psychology: A Lay Theories Approach to Self-Regulation, Social Perception, and Social Development. *American Psychologist*, 61(3), 192–203. <http://doi.org/10.1037/0003-066X.61.3.192>
- Moncrieff, J., Wessely, S., & Hardy, R. (1996). Active placebos versus antidepressants for depression. In *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd.
- Moore, G. E. (1960). *Principia Ethica:(1903)* (Vol. 960). Cambridge University Press Cambridge.
- Morgan, A. (2015). Is psychiatry dying? Crisis and critique in contemporary psychiatry. *Social Theory & Health*, 13(2), 141-161.
- Morton, T. A., Hornsey, M. J., & Postmes, T. (2009). Shifting ground: The variable use of essentialism in contexts of inclusion and exclusion. *British Journal of Social Psychology*, 48(1), 35–59. <http://doi.org/http://dx.doi.org/10.1348/014466607X270287>
- Morton, T. A., & Postmes, T. (2009). When differences become essential: Minority essentialism in response to majority treatment. *Personality and Social Psychology Bulletin*, 35(5), 656–668. <http://doi.org/http://dx.doi.org/10.1177/0146167208331254>
- Morton, T. A., Postmes, T., Haslam, S., & Hornsey, M. J. (2009). Theorizing gender in the face of social change: Is there anything essential about essentialism? *Journal of Personality and Social Psychology*, 96(3), 653–664. <http://doi.org/http://dx.doi.org/10.1037/a0012966>
- Most, T., Wiesel, A., & Blitzer, T. (2007). Identity and attitudes towards cochlear implant among deaf and hard of hearing adolescents. *Deafness & Education International*, 9(2), 68–82. <http://doi.org/10.1002/dei.207>
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children’s

- motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33.
- Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of personality and social psychology*, 89(6), 852.
- Munafò, M. R., Zammit, S., & Flint, J. (2014). Practitioner Review: A critical perspective on gene-environment interaction models - what impact should they have on clinical perceptions and practice? *Journal of Child Psychology and Psychiatry*, 55(10), 1092–1101. <http://doi.org/10.1111/jcpp.12261>
- Murali, V., & Oyebode, F. (2004). Poverty, social inequality and mental health. *Advances in psychiatric treatment*, 10(3), 216-224.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92(3), 289.
- Murphy, D., & Woolfolk, R. L. (2000). The harmful dysfunction analysis of mental disorder. *Philosophy, Psychiatry, & Psychology*, 7(4), 241-252.
- Murray, D., & Stoessl, A. J. (2013). Mechanisms and therapeutic implications of the placebo effect in neurological and psychiatric conditions. *Pharmacology & Therapeutics*, 140(3), 306–318. <http://doi.org/10.1016/j.pharmthera.2013.07.009>
- Nasser, M. (1987). Psychiatry in Ancient Egypt. *Psychiatric Bulletin*, 11(12), 420–422. <http://doi.org/10.1192/pb.11.12.420>
- Neale, B. M., Kou, Y., Liu, L., Ma'Ayan, A., Samocha, K. E., Sabo, A., ... others. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*, 485(7397), 242–245.
- Oakes, P. J. (1987). The salience of social categories. *Rediscovering the Social Group: A Self-Categorization Theory*, 117–141.
- Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members:

- The role of fit in the salience of social categorizations. *British Journal of Social Psychology*, 30(2), 125-144.
- Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W., & Laland, K. N. (2013). Niche Construction Theory: A Practical Guide for Ecologists. *The Quarterly Review of Biology*, 88(1), 3-28. <http://doi.org/10.1086/669266>
- Onorato, R. S., & Turner, J. C. (2004). Fluidity in the self-concept: the shift from personal to social identity. *European Journal of Social Psychology*, 34(3), 257-278.
- Osterman, J. E., & De Jong, J. (2007). Cultural issues and trauma. *Handbook of PTSD: Science and Practice*, 425-446.
- Parsons, T. (1951). Illness and the role of the physician: a sociological perspective. *American Journal of Orthopsychiatry*, 21, 452-460.
- Pemberton, R., & Wainwright, T. (2014a). The end of mental illness thinking? *International Journal of Clinical and Health Psychology*, 14(3), 216-220. <http://doi.org/10.1016/j.ijchp.2014.05.003>
- Peterson, C., & Seligman, M. E. (1984). Causal explanations as a risk factor for depression: theory and evidence. *Psychological Review*, 91(3), 347.
- Petersen, S., van den Berg, R. A., Janssens, T., & Van den Bergh, O. (2011). Illness and symptom perception: A theoretical approach towards an integrative measurement model. *Clinical Psychology Review*, 31(3), 428-439.
- Petrie, K. J., Weinman, J., Sharpe, N., & Buckley, J. (1996). Role of patients' view of their illness in predicting return to work and functioning after myocardial infarction: longitudinal study. *BMJ*, 312(7040), 1191-1194.
- Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression. *The British Journal of Psychiatry*, 177(6), 486-492. <http://doi.org/10.1192/bjp.177.6.486>

- Pilgrim, D. (2002). The biopsychosocial model in Anglo-American psychiatry: Past, present and future? *Journal of Mental Health, 11*(6), 585–594.
<http://doi.org/10.1080/09638230020023930>
- Pilgrim, D. (2013). The Failure of Diagnostic Psychiatry and some prospects of Scientific Progress Offered by Critical Realism. *Journal of Critical Realism, 12*(3), 336–358.
<http://doi.org/10.1179/1476743013Z.0000000004>
- Pingani, L., Luciano, M., Sampogna, G., Rosa, C. D., Pinna, F., Volpe, U., ... Fiorillo, A. (2014). The crisis in psychiatry: A public health perspective. *International Review of Psychiatry, 26*(4), 530–534. <http://doi.org/10.3109/09540261.2014.931838>
- Piper, A., & Merskey, H. (2004). The persistence of folly: A critical examination of dissociative identity disorder. Part I. The excesses of an improbable concept. *Canadian Journal of Psychiatry, 49*, 592–600.
- Platow, M. J., van Knippenberg, D., Haslam, S. A., van Knippenberg, B., & Spears, R. (2006). A special gift we bestow on you for being representative of us: Considering leader charisma from a self-categorization perspective. *British Journal of Social Psychology, 45*(2), 303–320. <http://doi.org/10.1348/014466605X41986>
- Platow, M. J., Hunter, J. A., Branscombe, N. R., & Grace, D. M. (2014). Social creativity in Olympic medal counts: Observing the expression of ethnocentric fairness. *Social Justice Research, 27*(3), 283–304.
- Platow, M. J., Voudouris, N. J., Coulson, M., Gilford, N., Jamieson, R., Najdovski, L., ... Terry, L. (2007). In-group reassurance in a pain setting produces lower levels of physiological arousal: Direct support for a self-categorization analysis of social influence. *European Journal of Social Psychology, 37*(4), 649–660.
<http://doi.org/http://dx.doi.org/10.1002/ejsp.381>

- Plomin, R., & Crabbe, J. (2000). DNA. *Psychological Bulletin*, 126(6), 806–828.
<http://doi.org/10.1037//0033-2909.126.6.806>
- Pöyhönen, S. (2014). Explanatory power of extended cognition. *Philosophical Psychology*, 27(5), 735-759.
- Porter, D., & Zachar, P. (2012). Recovery and the partitioning of scientific authority in psychiatry. *Recovery of People with Mental Illness: Philosophical and Related Perspectives*, 203.
- Postmes, T. (2003). A social identity approach to communication in organizations. *Social identity at work: Developing theory for organizational practice*, 81, 191.
- Postmes, T., Haslam, S. A., & Jans, L. (2012). A single-item measure of social identification: Reliability, validity, and utility. *The British Journal of Social Psychology / the British Psychological Society*. <http://doi.org/10.1111/bjso.12006>
- Postmes, T., Spears, R., & Lea, M. (1998). Breaching or Building Social Boundaries? SIDE-Effects of Computer-Mediated Communication. *Communication Research*, 25(6), 689–715. <http://doi.org/10.1177/009365098025006006>
- Prah, P., Copas, A. J., Mercer, C. H., Clifton, S., Erens, B., Phelps, A., ... & Johnson, A. M. (2013). Consistency in reporting sensitive sexual behaviours in Britain: change in reporting bias in the second and third National Surveys of Sexual Attitudes and Lifestyles (Natsal-2 and Natsal-3). *Sexually transmitted infections*.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2003). *Simple intercepts, simple slopes, and regions of significance in MLR 2-way interactions*. Retrieved July, 24, 2005.
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, 42(1), 185-227.

- Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47(9), 1102.
- Pull, C. B. (2012). Current status of knowledge on public-speaking anxiety. *Current opinion in psychiatry*, 25(1), 32-38.
- Putnam, H. (1980). Philosophy and our mental life. *Readings in the Philosophy of Psychology*, 1, 134–143.
- Quinn, D. M., Kahng, S. K., & Crocker, J. (2004). Discreditable: Stigma effects of revealing a mental illness history on test performance. *Personality and Social Psychology Bulletin*, 30(7), 803–815. <http://doi.org/http://dx.doi.org/10.1177/0146167204264088>
- Rangel, U., & Keller, J. (2011). Essentialism goes social: belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology*, 100(6), 1056.
- Rapee, R. M. (1998). *Overcoming shyness and social phobia: A step-by-step guide*. Jason Aronson.
- Reddy, V. P., Kozielska, M., Johnson, M., Vermeulen, A., de Greef, R., Liu, J., ... Proost, J. H. (2011). Structural models describing placebo treatment effects in schizophrenia and other neuropsychiatric disorders. *Clinical Pharmacokinetics*, 50(7), 429–450.
- Redersdorff, S., & Guimond, S. (2006). Comparing oneself over time: The temporal dimension in social comparison. In *Social comparison and social psychology: Understanding cognition, intergroup relations, and culture* (pp. 76–96). New York, NY: Cambridge University Press; US.
- Reicher, S. D., Spears, R., & Postmes, T. (1995). A Social Identity Model of Deindividuation Phenomena. *European Review of Social Psychology*, 6(1), 161–198. <http://doi.org/10.1080/14792779443000049>

- Reicher, S., Hopkins, N., Levine, M., & Rath, R. (2005). Entrepreneurs of hate and entrepreneurs of solidarity: Social identity as a basis for mass communication. *International Review of the Red Cross*, 87(860), 621–637.
<http://doi.org/10.1017/S1816383100184462>
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Social Science & Medicine*, 90, 24–31.
- Remington, G., & Shammi, C. (2005). Overstating the case about recovery?. *Psychiatric Services*, 56(8), 1022–1022.
- Resnick, S. G., Fontana, A., Lehman, A. F., & Rosenheck, R. A. (2005). An empirical conceptualization of the recovery orientation. *Schizophrenia Research*, 75(1), 119–128. <http://doi.org/10.1016/j.schres.2004.05.009>
- Reynolds, K. J., Oakes, P. J., Haslam, S. A., Nolan, M. A., & Dolnik, L. (2000). Responses to powerlessness: Stereotyping as an instrument of social conflict. *Group dynamics: Theory, research, and practice*, 4(4), 275.
- Riley, S., Rodham, K., & Gavin, J. (2009). Doing weight: Pro-ana and recovery identities in cyberspace. *Journal of Community & Applied Social Psychology*, 19(5), 348–359.
<http://doi.org/10.1002/casp.1022>
- Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., ... Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*, 45(10), 1150–1159.
<http://doi.org/10.1038/ng.2742>
- Rips, L. J. (1989). Similarity and Analogical Reasoning. In S. Vosniadou & A. Ortony (Eds.) (pp. 21–59). New York, NY, USA: Cambridge University Press. Retrieved from <http://dl.acm.org/citation.cfm?id=107328.107330>

- Ritsher, J. B., Otilingam, P. G., & Grajales, M. (2003). Internalized stigma of mental illness: psychometric properties of a new measure. *Psychiatry Research, 121*(1), 31–49.
- Robert, B. (2012). *The Anatomy of Melancholy-The Original Classic Edition*. Emereo Publishing.
- Roe, D., Rudnick, A., & Gill, K. J. (2007). The concept of "being in recovery". *Psychiatric Rehabilitation Journal, 30*(3), 171.
- Rosenfield, S. (1997). Labeling Mental Illness: The Effects of Received Services and Perceived Stigma on Life Satisfaction. *American Sociological Review, 62*(4), 660.
<http://doi.org/10.2307/2657432>
- Rosenquist, J. N., Fowler, J. H., & Christakis, N. A. (2011). Social network determinants of depression. *Molecular Psychiatry, 16*(3), 273–281. <http://doi.org/10.1038/mp.2010.13>
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds?. Semin, Gün R. (Ed); Fiedler, Klaus (Ed), *Language, interaction and social cognition*. (pp. 11-36). Thousand Oaks.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied, 80*(1), 1.
- Rudnick, A. (2008). Recovery from Schizophrenia: A Philosophical Framework. *American Journal of Psychiatric Rehabilitation, 11*(3), 267–278.
<http://doi.org/10.1080/15487760802186360>
- Sadler, J. Z., & Agich, G. J. (1995). Diseases, Functions, Values, and Psychiatric Classification. *Philosophy, Psychiatry, & Psychology, 2*(3), 219–231.
- Sadler, J. Z., Wiggins, O. P., & Schwartz, M. A. (1994). *Philosophical perspectives on psychiatric diagnostic classification*. Johns Hopkins University Press. Retrieved from <http://psycnet.apa.org.virtual.anu.edu.au/psycinfo/1994-97760-000>

- Saguy, A. C., & Ward, A. (2011). Coming out as fat rethinking stigma. *Social Psychology Quarterly*, 74(1), 53-75.
- Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A Systematic Review of the Prevalence of Schizophrenia. *PLoS Med*, 2(5), e141.
<http://doi.org/10.1371/journal.pmed.0020141>
- Sass, H. (2009). *Anthology of German Psychiatric Texts*. John Wiley & Sons.
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379.
- Scheff, T. J. (1963). The Role of the Mentally Ill and the Dynamics of Mental Disorder: A Research Framework. *Sociometry*, 26(4), 436-453. <http://doi.org/10.2307/2786147>
- Scheff, T. J. (1970). *Being mentally ill: A sociological theory*. Transaction Publishers.
- Scheff, T. J. (1974). The Labelling Theory of Mental Illness. *American Sociological Review*, 39(3), 444. <http://doi.org/10.2307/2094300>
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of personality and social psychology*, 85(3), 440.
- Schrader, S., Jones, N., & Shattell, M. (2013). Mad pride: Reflections on sociopolitical identity and mental diversity in the context of culturally competent psychiatric care. *Issues in Mental Health Nursing*, 34(1), 62-64.
- Schroder, H. S., Dawood, S., Yalch, M. M., Donnellan, M. B., & Moser, J. S. (2016). Evaluating the Domain Specificity of Mental Health-Related Mind-Sets. *Social Psychological and Personality Science*.
- Schultz, W. (2015). Neuroessentialism Theoretical and Clinical Considerations. *Journal of Humanistic Psychology*.
- Schulze, B., & Angermeyer, M. C. (2003). Subjective experiences of stigma. A focus group

- study of schizophrenic patients, their relatives and mental health professionals. *Social Science & Medicine*, 56(2), 299–312.
- Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, 57(2), 149-174.
- Searle, J. R. (1995). *The construction of social reality*. Simon and Schuster.
- Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11(2), 107-130.
- Shih, M., Pittinsky, T. L., & Trahan, A. (2006). Domain-specific effects of stereotypes on performance. *Self and Identity*, 5(1), 1–14.
- Sigal, L. H., & Hassett, A. L. (2002). Contributions of societal and geographical environments to "chronic Lyme disease": the psychopathogenesis and aporology of a new "medically unexplained symptoms" syndrome. *Environmental health perspectives*, 110(Suppl 4), 607.
- Smith, L. G., & Postmes, T. (2011). Shaping stereotypical behaviour through the discussion of social stereotypes. *British Journal of Social Psychology*, 50(1), 74–98.
<http://doi.org/http://dx.doi.org/10.1348/014466610X500340>
- Smyth, L., Mavor, K. I., Platow, M. J., Grace, D. M., & Reynolds, K. J. (2013). Discipline social identification, study norms and learning approach in university students. *Educational Psychology*, 0(0), 1–21. <http://doi.org/10.1080/01443410.2013.822962>
- Solomon, A. (2012). *Far from the tree: Parents, children and the search for identity*. Simon and Schuster.
- Speak, G. (1990). An odd kind of melancholy: reflections on the glass delusion in Europe

(1440-1680. *History of Psychiatry*, 1(2), 191–206.

Spears, R., Postmes, T., Lea, M., & Wolbert, A. (2002). When are net effects gross products? The power of influence and the influence of power in computer-mediated communication. *Journal of Social Issues*, 58(1), 91–107.

<http://doi.org/http://dx.doi.org/10.1111/1540-4560.00250>

St Claire, L., & He, Y. (2009). How Do I Know if I Need a Hearing Aid? Further Support for the Self-Categorisation Approach to Symptom Perception. *Applied Psychology*, 58(1), 24–41. <http://doi.org/10.1111/j.1464-0597.2008.00380.x>

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *International Journal of Epidemiology*.

<http://doi.org/10.1093/ije/dyu038>

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. <http://doi.org/http://dx.doi.org/10.1037/0022-3514.69.5.797>

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 34, pp. 379–440). Academic Press. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0065260102800090>

Stier, M. (2013). Normative preconditions for the assessment of mental disorder. *Frontiers in Psychology*, 4. <http://doi.org/10.3389/fpsyg.2013.00611>

Storms, M. D., & McCaul, K. D. (1976). Attribution processes and emotional exacerbation

- of dysfunctional behavior. *New Directions in Attribution Research*, 1, 143-164.
- Stryker, S., & Statham, A. (1985). Symbolic interaction and role theory. Retrieved from <http://www.citeulike.org/group/940/article/693984>
- Szasz, T. (1961). *Myth of Mental Illness* (Vol. 15). New York.
- Taiminen, T. J., Kallio-Soukainen, K., Nokso-Koivisto, H., Kaljonen, A., & Helenius, H. (1998). Contagion of deliberate self-harm among adolescent inpatients. *Journal of the American Academy of Child & Adolescent Psychiatry*, 37(2), 211–217.
- Henri, T. (1974). Social identity and intergroup behaviour. *Social Science Information*, 13(2), 65.
- Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.
- Tajfel, H. (1979). Individuals and groups in social psychology. *British Journal of Social and Clinical Psychology*, 18(2), 183-190.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations*, 33, 47.
- Tarrant, M., & Butler, K. (2011). Effects of self-categorization on orientation towards health. *British Journal of Social Psychology*, 50(1), 121–139.
<http://doi.org/10.1348/014466610x511645>
- Tarrant, M., Hagger, M. S., & Farrow, C. V. (2011). Promoting positive orientation towards health through social identity. *The social cure: Identity, health and well-being*, 39.
- Taylor, J., & Deane, F. P. (2002). Development of a short form of the Test Anxiety Inventory (TAI). *Journal of General Psychology*, 129(2), 127–136.
- Terry, D. J., & Hogg, M. A. (1996). Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification. *Personality and Social Psychology Bulletin*, 22(8),

776–793. <http://doi.org/10.1177/0146167296228002>

Terry, D. J., Hogg, M. A., & White, K. M. (1999). The theory of planned behaviour: Self-identity, social identity and group norms. *British Journal of Social Psychology*, 38(3), 225–244. <http://doi.org/10.1348/014466699164149>

Thoits, P. A. (1985). Self-Labeling Processes in Mental Illness: The Role of Emotional Deviance. *American Journal of Sociology*, 91(2), 221–249.

Timimi, S. (2014). No more psychiatric labels: Why formal psychiatric diagnostic systems should be abolished. *International Journal of Clinical and Health Psychology*, 14(3), 208–215. <http://doi.org/10.1016/j.ijchp.2014.03.004>

Titov, N., Andrews, G., Schwencke, G., Drobny, J., & Einstein, D. (2008). Shyness 1: distance treatment of social phobia over the Internet. *Australian and New Zealand Journal of Psychiatry*, 42(7), 585–594. <http://doi.org/10.1080/00048670802119762>

Turkheimer, E. (2011). Genetics and Human Agency: Comment on Dar-Nimrod and Heine (2011). *Psychological Bulletin*, 137(5), 825–828.

Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology*, 5(1), 1–34.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. *Advances in Group Processes*, 2, 77–122.

Turner, J. C. (1991). *Social influence*. Belmont, CA: Thomson Brooks/Cole Publishing Co; US.

Turner, J. C. (2010). The analysis of social influence. In *Rediscovering social identity* (pp. 273–286). New York, NY: Psychology Press; US.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987b). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Basil

Blackwell; US.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence.

British Journal of Social Psychology, 25(3), 237-252.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective:

Cognition and social context. *Personality and Social Psychology Bulletin*, 20, 454–454.

Turner, J. C., Wetherell, M. S., & Hogg, M. A. (1989). Referent informational influence and group polarization. *British Journal of Social Psychology*, 28(2), 135-147.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of educational research*, 78(4), 751-796.

Valenstein, E. S. (1986). *Great and desperate cures: The rise and decline of psychosurgery and other radical treatments for mental illness* (Vol. xiv). New York, NY, US: Basic Books.

Valentiner, D. P., Jencius, S., Jarek, E., Gier-Lonsway, S. L., & McGrath, P. B. (2013). Pre-treatment shyness mindset predicts less reduction of social anxiety during exposure therapy. *Journal of Anxiety Disorders*, 27(3), 267-271.

Valentiner, D. P., Mounts, N. S., Durik, A. M., & Gier-Lonsway, S. L. (2011). Shyness mindset: Applying mindset theory to the domain of inhibited social behavior.

Personality and Individual Differences, 50(8), 1174–1179.

<http://doi.org/10.1016/j.paid.2011.01.021>

Van Ommeren, M., de Jong, J. T., Sharma, B., Komproe, I., Thapa, S. B., & Cardena, E.

(2001). Psychiatric disorders among tortured Bhutanese refugees in Nepal. *Archives of General Psychiatry*, 58(5), 475-482.

- Vaughan, G. M. (1978). Social change and intergroup preferences in New Zealand. *European Journal of Social Psychology*, 8(3), 297–314.
<http://doi.org/10.1002/ejsp.2420080304>
- Vernon, D. T. (1974). Modeling and birth order in responses to painful stimuli. *Journal of Personality and Social Psychology*, 29(6), 794.
- von Bertalanffy, L. (1956). {General System Theory}. *General Systems*, 1, 1–10.
- Vredenburg, K., Flett, G. L., & Krames, L. (1993). Analogue versus clinical depression: A critical reappraisal. *Psychological Bulletin*, 113(2), 327–344.
<http://doi.org/10.1037/0033-2909.113.2.327>
- Wahl, O. F. (1999). Mental health consumers' experience of stigma. *Schizophrenia Bulletin*, 25(3), 467.
- Wakefield, J. C. (1999). Evolutionary versus prototype analyses of the concept of disorder. *Journal of Abnormal Psychology*, 108(3), 374.
- Wakefield, J. C. (2007). The concept of mental disorder: diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, 6(3), 149.
- Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression: variable, substantial, and growing. *JAMA*, 287(14), 1840-1847.
- Walter, H. (2013). The third wave of biological psychiatry. *Frontiers in Psychology*, 4, 582.
- Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress Contagion: Physiological Covariation Between Mothers and Infants. *Psychological Science*, 25(4), 934–942.
<http://doi.org/10.1177/0956797613518352>
- Wegner, D. M. (1997). When the antidote is the poison: Ironic mental control processes. *Psychological Science*, 8(3), 148-150.
- Wegner, D. M., Broome, A., & Blumberg, S. J. (1997). Ironic effects of trying to relax under

- stress. *Behaviour research and therapy*, 35(1), 11-21.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548–573. <http://doi.org/10.1037/0033-295X.92.4.548>
- Weisberg, H. (1980). *Statistical Methods for Comparative Studies*.
- Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H.-G., ... others. (1996). Cross-national epidemiology of major depression and bipolar disorder. *Jama*, 276(4), 293–299.
- Wessely, S. (1987). Mass hysteria: two syndromes? *Psychological Medicine*, 17(1), 109–120.
- White, P. D., Chalder, T., & Sharpe, M. (2014). The planning, implementation and publication of a complex intervention trial for chronic fatigue syndrome: the PACE trial. *Psychiatric Bulletin*, pb.bp.113.045005. <http://doi.org/10.1192/pb.bp.113.045005>
- Whitt, H. P., Meile, R. L., & Larson, L. M. (1979). Illness role theory, the labeling perspective and the social meanings of mental illness: an empirical test. *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology*, 13, 655–666.
- Wichers, M. (2014). The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, 44(7), 1349–1360. <http://doi.org/10.1017/S0033291713001979>
- Wild, C. P. (2005). Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology Biomarkers & Prevention*, 14(8), 1847–1850. <http://doi.org/10.1158/1055-9965.EPI-05-0456>
- Williams, N. E. (2011). Arthritis and Nature’s joints. *Carving Nature at Its Joints: Natural*

Kinds in Metaphysics and Science, 199–230.

Williams, S. R., Cash, E., Daup, M., Geronimi, E. M. C., Sephton, S. E., & Woodruff-

Borden, J. (2013). Exploring patterns in cortisol synchrony among anxious and nonanxious mother and child dyads: A preliminary study. *Biological Psychology*, 93(2), 287–295. <http://doi.org/10.1016/j.biopsycho.2013.02.015>

Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences-Cognition*. Cambridge University Press.

Yap, P. (1974). *Comparative psychiatry: A theoretical framework* (Vol. 3). Published for the Clarke Institute of Psychiatry by University of Toronto Press.

Young, J. E. (1994). *Cognitive therapy for personality disorders: A schema-focused approach*. Professional Resource Press/Professional Resource Exchange. Retrieved from <http://psycnet.apa.org.virtual.anu.edu.au/psycinfo/1994-98188-000>

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema Therapy: A Practitioner's Guide*. Guilford Press.

Yzerbyt, V., Castano, E., Leyens, J.-P., & Paladino, M.-P. (2000). The Primacy of the Ingroup: The Interplay of Entitativity and Identification. *European Review of Social Psychology*, 11(1), 257–295. <http://doi.org/10.1080/14792772043000059>

Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The Interplay of Subjective Essentialism and Entitativity in the Formation of Stereotypes. *Personality and Social Psychology Review*, 5(2), 141–155. http://doi.org/10.1207/S15327957PSPR0502_5

Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology*, 40(3), 424–431.

<http://doi.org/10.1016/j.jesp.2003.10.001>

Zachar, P. (2000). Psychiatric Disorders Are Not Natural Kinds. *Philosophy, Psychiatry, & Psychology*, 7(3), 167–182.

Zachar, P. D. & Kendler, K. (2007). Psychiatric Disorders: A Conceptual Taxonomy. *American Journal of Psychiatry*, 164(4), 557–565.

<http://doi.org/10.1176/appi.ajp.164.4.557>

Zhang, W., Robertson, J., Jones, A. C., Dieppe, P. A., & Doherty, M. (2008). The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Annals of the Rheumatic Diseases*, 67(12), 1716–1723.

<http://doi.org/10.1136/ard.2008.092015>

Zimbardo, P. G., Pilkonis, P. A., & Norwood, R. M. (1975). SOCIAL DISEASE CALLED SHYNESS. *Psychology Today*, 8(12), 69.

Zimbardo, P. G. (1977). *Shyness: What it is, what to do about it* Addison-Wesley. Reading, MA.

Zucker, K. J., & Lawrence, A. A. (2009). Epidemiology of gender identity disorder: Recommendations for the standards of care of The World Professional Association for Transgender Health. *International Journal of Transgenderism*, 11(1), 8-18.